RF Project 1465
Report 2

THE OHIO STATE UNIVERSITY

RESEARCH FOUNDATION

1314 KINNEAR ROAD COLUMBUS 12, OHIO

Project on Linguistic Analysis
Report No. 6

RESEARCH ON MANDARIN PHONOLOGY

W. S-Y. Wang, Division of Linguistics
K. P. Li, Department of Electrical Engineering
R. L. Brotzman, Department of Electrical Engineering

DEPARTMENT OF THE NAVY
Office of Naval Research
Contract No. Nonr-495(27)

GENERAL OBJECTIVES OF THE PROJECT ON LINGUISTIC ANALYSIS

The central objective of this group is to uncover and formalize some of the underlying properties of natural language. While our attention is primarily focused on Mandarin Chinese and American English, the results we achieve from studying these two languages will be of value toward the construction of a general theory of language. These results will mostly take the form of assertions regarding the structure of the object language under study (rules in the grammar) and statements regarding the structure of the metalanguage used to make these assertions (rules in the theory of language). Clearly these two types of results are mutually dependent. An example of a set of grammar rules is the paper on Mandarin syntax. An example of a set of metalinguistic statements is Fillmore's paper on "traffic rules" in a grammar. Both of these papers are in POLA No. 3.

A complete grammar of a language will contain a set of rules which will generate all and only the sentences of the language with their corresponding structural descriptions. This is an ideal toward which we strive in our analysis of Mandarin and English. In the process of grammar construction, we aim to formulate rules that are precise enough to be programable onto a computer. Such precision is necessary for making explicit the metalinguistic assumptions which underlie the grammar and for the latter to be evaluable for overall economy and consistency. When properly formulated, these rules should correctly describe such notions as grammatical ambiguity and relationship among sentences.

For the construction of a comprehensive theory of language that can explicate the ability to use language, i.e., to analyze, produce, understand and paraphrase sentences, extensive knowledge in the area of semantic analysis is needed to support the grammatical knowledge. In order to effectively understand our ability to translate between languages, further transfer rules would need to be developed in terms of which relations can be established among the grammatic-semantic structures of different languages. We feel that any successful

simulation of linguistic processes, such as sentence recognition and translation, will ultimately need to be based on such a comprehensive theory of language.

The kind of research that we have been pursuing will provide a grammatical basis toward the construction of such a theory. Although we have not yet actively extended our work into semantic analysis and transfer rules, our present research will lead into it in a natural way at the appropriate stage in its development. At that time, the close association that has been maintained between the development of the two grammars will prove to be especially valuable in discovering the transfer rules.

A part of our overall effort has been directed to the improvement of communication and cooperation between ourselves and other groups working on related problems. There exists a healthy situation in the case of people working on English grammar and linguistic theory from a rigorous point of view, and interchange of information is free and effective.

The picture is very different with work in Chinese grammar. While there is massive and vigorous activity on Chinese grammar and linguistic theory in the Communist countries, especially in China, a disproportionately small amount of this work is available in this country. In addition to the political barriers which cause delays and miscarriages in the mail, among other inconveniences, there is an equally serious obstacle even with the available materials. This is the problem of understanding each other's work through the maze of different linguistic frameworks and terminologies. Here the mutual intelligibility of the various dialects of Linguistics is very low. Nevertheless, we feel that we can learn much from the activities abroad which, at least in the area of Chinese grammar, exist on a much grander scale than in this country.

To create better channels of technical communication, we have compiled a bibliography with glossary, POLA No. 5. Some of the translations we have made from Chinese and Japanese into English appear in POLA No. 4. POLA No. 6 contains tables and algorithms for converting among the various systems of Mandarin transcription.

We are also issuing a book in Chinese, entitled Biànhuànlǜ Yǔfǎ Lǐlùn, which contains the theoretical framework of our research. This book is in large part based on Chomsky's well-known Syntactic Structures.

On the next page is a list of technical reports which have been or will be sent out shortly. The purpose of distributing these reports is twofold. One is that it gives us a quick method of making known our results to a restricted audience without the necessary delay which goes with journal publication. These reports also permit the distribution of materials which are not suitable for inclusion in a journal article, such as large amounts of data. The other purpose is that we will be enabled to receive comments and criticisms on our work, and perhaps incorporate these therein, before submitting some of the results for publication in open literature.

W. S-Y. Wang
March 22, 1963

POLA No. 1: C. J. Fillmore, Indirect object constructions in English and the ordering of transformations; 49pp, February 1962.

POLA No. 2: W. S-Y. Wang, Synchronic studies in Mandarin grammar: a selected bibliography; 15pp, February 1962.

POLA No. 3: C. J. Fillmore, The position of embedding transformations in a grammar;* W. S-Y. Wang, Some syntactic rules in Mandarin;** L. F. Meyers and W. S-Y. Wang, Tree representations in linguistics; 113pp., May 1963.

POLA No. 4: Translations of works on Chinese grammar: The phonemes of the Peiping dialect; Adjectivals and adverbials; Ba sentences and bei sentences; 104pp, May 1963.

POLA No. 5: W. S-Y. Wang and L. Liu, Bibliography and glossary for Chinese grammar.

POLA No. 6: W. S-Y. Wang, Mandarin phonology; R. Brotzman, Vowel formant values; K. P. Li, Tone perception experiment with appended test materials; Inventory of Mandarin syllables; Syllable inventory according to the Pinyin alphabet; Conversion table for different transcriptions; Conversion algorithms among different transcriptions.

POLA No. 7: E. Ching, Four syllable phrases in Chinese; S. Annear and L. Liu, A fragment of Mandarin grammar: Ba and bei; A. T. Tsai, Resultative verb in Mandarin; C. J. Fillmore, Complement structures in English.

*to appear in the journal Word.

**to appear in the Proceedings of the Ninth International Congress of Linguists.

MANDARIN PHONOLOGY*

W. S-Y. Wang

This paper is a discussion of research now being pursued on Mandarin phonology at the Ohio State University. The discussion will be in three parts. First, some comments will be given on the theoretical framework which underlies our efforts. Secondly, I will report on the methods we have used in taking accoustical measurements on the syllable-finals and our accomplishments so far. Thirdly, the results will be presented of a psycholinguistic experiment that was recently performed to determine the phonemic distinctness of certain tonal phenomena. It will become evident that these three parts of the present discussion are all oriented toward one central goal--the construction of the simplest phonology for Mandarin.

As is well-known, the concept of simplicity houses many variables which are ill-defined and subjective. Clearly, we do not want our phonology to be simple at the expense of making the rest of the grammar complicated. For our present purposes, let us assume, falsely to be sure, that the syntax will not be affected by the decisions made in the phonology. The simplest phonology then will be the simplest conjunction of (1) the base forms of all the morphemes in some economical representation, and (2) the phonological rules which can convert all grammatical concatenations of these morphemes in their base form into the correct physical signals of speech.

Although the state of knowledge in phonological studies does not permit the making of a definite claim, it seems that such a phonology is really a partial explanatory model for our ability to speak. Here I am limiting this ability to "how to pronounce a given expression in the language." That is to say, the phonology should be able to convert (1) which is in a morphophonemic representation into (2) which is in a narrow phonetic representation.

Syllable: 1 2 3 4 5 6 7 8 9
(1) Zhèi-ge "lin"-zi, shi Lǎo Lǐ xiě de.
(2)#dʐ̹è^{i}gəlíndzʂ̩lǎ^{u}lǐç̩^{i}ʑ̌də#

*A version of this paper was read before a section meeting of the Modern Languages Association, December 1962, Washington, D. C.

Some of the matters that the phonology would need to take care of in this example are the following: The last sound in syllable 1 is a glide toward the high-front position whose amplitude is lower than that for vowels. Although the consonant in syllable 2 is usually unvoiced, in the present context, that is, intervocalic and unstressed, it becomes voiced. In syllable 3, the vowel is lax, i.e., "I", whereas it would be tense if it were not followed by a nasal, cf. the one in syllable 7. Furthermore, we know that since the nasal is acute rather than grave it has about the same duration as the preceding noncompact vowel. Where the nasal is grave, say as in "ling", it is usually two or three times longer than the preceding vowel. The high vowels in syllables 4 and 5 may be deleted. The tone 3 in Lao changes into a tone 2 because it is followed by another tone 3 in the same syntactic construction. As we shall see later, this derived tone 2 is not distinguishable from the regular tone 2 under normal circumstances. Syllable 7 may also change into tone 2 in more casual forms of speech.

The extent to which we could narrow down on the notation in (2) depends of course on the amount of physical information we have available. The more narrow, or the more precise, our phonetic notation, the closer we come in linking these marks on paper to the real world of speech sounds. And until we are able to perform this last step, say with a speech synthesizer, the phonology will not have been validated in the strictest sense. This is the motivation for enlarging our fund of knowledge about the basic physical properties of Mandarin speech.

On the other hand, the specification of the base form of the morphemes ought to be as simple as possible. It is well-known that man is quite limited in his ability to process information. This limitation is in terms of both discrimination among and remembrance of signals. Yet the amount of physical information contained in the speech wave that he produces is extremely high. If we quantize the wave into 128 levels and sample it 8000 times per second, which is equivalent to the capacity of a channel of bandwidth 5 kc and a signal-to-noise ratio of 30 db, then the information flows at 56000 bits per second.

This is clearly much too high for the speaker to be able to deliberately send, or for the hearer to process effectively. In a sense, part of the task of the phonologist is to sort all this information into three kinds: (1) the kind that is predictable from knowledge about the human speech mechanism, (2) the kind that is predictable from knowledge about the language, and (3) the kind

that is not predictable. (1) includes those properties which I have elsewhere called intrinsic, and form a subset of the phonological universals of language. (2) includes those properties, called extrinsic, which the particular language assigns to the speech wave. To a large extent, what we learn is (2) when we learn the "pronunciation" of a particular language.

In order to reduce the amount of information contained in the base forms, and thereby make the phonology a more plausible model in view of the human limitations in handling information, we must make sure that only information of the third kind is specified. Since (1) and (2) are not based on any decisions made by the speaker (nor do they necessitate prolonged attention by the hearer), they may be introduced by the phonological rules. Thus, according to this model, the only phonological information that the speaker and/or the hearer need to process is the third type.

Assuming an average transmission of six syllables per second with fifteen bits of information per syllable, the information rate is only 90 bits per second. This is not too unreasonable a figure since the language is simultaneously redundant at many levels.

This approach also turns out to agree quite well with a rather straightforward conception of simplicity, namely, ceteris paribus, the grammar with fewer restrictions (or shorter grammar) is simpler. If many morphemes all obey a certain restriction, clearly we would obtain a simpler grammar by stating this restriction just once in a rule rather than by repeating it in the base form of each of these morphemes.

One of the beautiful aspects of the phonological systems of natural languages is the correlation between distributional classes and phonetic classes of sounds. This means that many classes of sounds, for which we want to write phonological rules, will share a set of phonetic features that is not shared by other sounds. Furthermore, it is more economical to specify such sounds as a class than to specify any individual sound in the class. This is necessarily so since for the latter we have to additionally enumerate those features that this sound does not share with the phonetic class of which it is a member.

These two observations lead us to conclude that we should use some collection of phonetic features as our basic unit both in the coding of the base forms of the morphemes and in the statement of the phonological rules. The notion of "phonetic feature" has existed since antiquity, and there are many collections

available which are intended to be exhaustive for use on natural languages. For our research, we have accepted the collection of distinctive features proposed by Jakobson and his associates.

Part of our present research on Mandarin phonology consists of searching for the most economical coding of the base forms of the morphemes and the associated rules. Our experience so far is that for an average syllable, only some fifteen distinctive features need to be specified in the base form. In an average fully specified syllable, the number of features mentioned is around forty. The conversion from the minimally required features to the full specification of features in the syllable is the function of phonological rules. These rules will probably include the following types:

(1) RS rules--those which add the retroflex suffix and effect a change on the final of the syllable;

(2) SV rules--those which change vowels in certain positions to glides;

(3) RD rules--those which reduplicate syllables;

(4) SA rules--those which assign stress;

(5) TS rules--those which provide for the correct tone sandhi;

(6) TP rules--those which place the tone on the proper portion of the syllable.

Although this part of our research is far from being completed, it is clear that for the purposes of simplicity, certain groups of rules as well as certain rules within groups need to be ordered in their application. For instance, the tonal value of a syllable is marked at the front end of the syllable in the early part of the phonology. Some of the reasons for doing this are as follows: We can save a feature in specifying the high rising tone for syllables beginning with m, n, l, and r since these syllables do not occur with the high level tone; all we need to do is specify the highness of the tone and put in the "nonlevelness" by rule. More importantly, placing the tone mark at the beginning of the syllable greatly facilitates the rule for reduplication. Consequently, the rules for tone placement must apply after the two above considerations have been taken care of.

Since this kind of phonology is stated precisely, its rules are easily amenable to being mechanized. We hope to program these rules, after they are completed, onto a computer for overall validation.

After the application of some of the phonological rules, when the syllables are all spelled out in fully specified segments, the job is about half done. It remains to find out what are the physical properties, in terms of frequency, amplitude, and time, that are represented by these features. If we connected a synthesizer with the computer program mentioned above, then the phonology would be complete with hardware. The input would be the incompletely specified base form representations of the morphemes and the output would be Mandarin speech.

To accomplish the latter half of this job would of course require a great deal of physical information about Mandarin speech. Since the birth of acoustic phonetics in the middle 1940's, much knowledge has been gathered about the physical nature of English, Russian, French, Japanese and Swedish. Unfortunately, such information is severely lacking for Mandarin. In starting out to gather such information, we have begun with the citation syllable in a fixed frame. There are 1192 such syllables. The frame sentence we have chosen is the following:

(3) Zhèi-ge X-zi shi Lǎo Lǐ xiěde,

where X occupies the position of the citation syllable. Six informants of the Peking dialect (3 men and 3 women) read these onto tape under controlled acoustical conditions. This represents a total of some 15 hours of tape recording.

The frame was used for the obvious reason of maintaining constant the gross properties of speech, such as pitch, loudness and rhythm.

To date, the recordings of one of the informants have been completely measured with respect to the following parameters for the citation syllable: the frequencies of the first three formants and the durations of the syllable, the syllable final, and the sentence. The formant frequencies were estimated from broad-band sonograms and narrow-band amplitude sections. We have tried to take at least one set of formant measurements for each vowel in the syllable. Thus for one informant, we have taken some 10,000 measurements for the parameters mentioned. Some of these measurements are presented in Appendix I of "Vowel Formant Values" on p.12 of this report.

It is of course quite difficult to master the significance of all these numbers until after we have properly coded them for easy manipulation. We might mention one observation that is quite striking that pertains to the durational ratio of the vowel and the nasal in the syllable final. When the vowel is followed by an alveolar nasal, it is about as long as the nasal. However, if the

following nasal is velar, then the vowel may be only one-fourth as long. This property is definitely part of the extrinsic system of Mandarin phonology. In American English, for example, the opposite is reported to be true, although the ratio does not differ nearly as much. This observation ties in very well with the phonological rules of retroflexion: while the alveolar nasal completely disappears when followed by the retroflex suffix, the velar one does so only after nasalizing the preceding vowel.

Any research on Mandarin phonology has to take into account the complexities of the voice pitch. We are at present investigating the feasibility of various schemes of pitch extraction in order to avoid the great demands on time were we to do it with the Sonograph. It is clear however that on a gross level, we need to ascertain those pitch or tonal phenomena in the language which can be consistently distinguished. In this connection, we have conducted a small scale psycholinguistic experiment to investigate the question of whether tone 2 and tone 3 are consistently distinguishable from each other when they occur before another tone 3. A detailed description of this experiment can be found elsewhere in this report. (See "Tone Perception Experiment", pp. 19-26)

We can now say with a good deal of confidence that we cannot consistently distinguish tone 2 from tone 3 in this environment. No one has scored better than 62% correct, and this includes the informants themselves. Furthermore, there is little correlation between these scores with either the initial or the final of the test syllable. This last finding is somewhat disenchanting since we had expected a good correlation with the finals, especially with the mid voweld. A good positive correlation was found, however, between identifiability and meaningfullness.

VOWEL FORMANT VALUES

R. L. Bretzman

Objective:

The objective of this research is to determine the steady state values of the first three formants of Mandarin Chinese vowels. Inasmuch as spectrographic analysis was used to gather these data, other information is readily available should it become desirable to investigate additional properties of the acoustic signals involved.

Introduction:

Throughout this report, the symbols used will be those assigned by Lawton M. Hartman, III[1] unless specified otherwise. In appendices 1 and 2 the first symbol group of each set of data is in the Hartman notation with the target vowel underlined. Beneath this, in parentheses, is listed the IPA symbol for the target vowel felt to be most appropriate by W. S-Y. Wang and the author.

Procedure:

The 1192 citation syllables of Mandarin Chinese were assembled and arranged in a matrix of initial phoneme rows and vowel columns. A listing of these 1192 syllables is given on pp. 27-33. The structure of the citation syllable is illustrated in Fig. #1.

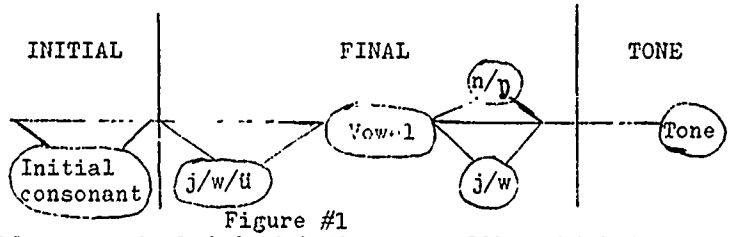


Figure #1

No syllables were included with the -er suffix, which is very common in Mandarin, thereby eliminating all nasalized vowels of the language. These syllables were read from characters one time each by six speakers of Mandarin Chinese. The linguistic background of each of the six informants is as follows:

[1] "The Segmental Phonemes of the Peiping Dialect", published in Readings in Linguistics, 1958, Martin Joos, ed.

Informant	Sex	Age	Chinese Dialect	Place of Birth	Elem. School	High School	College	Dialect spoken at home	Proficiency in English
#1	M	34	Mandarin	Peking	Peking	Peking	Peking/USA	Mandarin	Fluent
#2	F	34	Mandarin	Peking	Peking	Peking	Peking/USA	Mandarin	Fluent
#3	M	29	Mandarin	Peking	Peking	Taiwan	Taiwan	Mandarin	Fluent
#4	F	50	Mandarin	Peking	Peking	Peking	Nanking	Mandarin	Good
#5	F	45	Mandarin	Peking	Tientsin	Peking	Peking	Mandarin	Good
#6	M	32	Mandarin	Peking	Shanghai/ Peking	Shanghai/ Chungking	Nanking/ Taichung	Mandarin	Poor

Each syllable was read in the context of the carrier sentence "Zhèige ____ zì shì lǎo Lǐ xiě de" (Pinyin system), which means, "This word ____ was written by Mr. Li." The recordings were made in a sound-proof recording booth, and, in all cases except for informant #5 were made in a single session for each individual, the session lasting approximately 1 1/2 hours with a 10-15 minute rest midway during the session. "Scotch 175" tape was used in all recordings and was recorded full track, at 7 1/2 i.p.s. on an Ampex 351 recorder using an RCA 77-DX ribbon microphone with all informants except informant #3 who used an E.V. 666 dynamic microphone.

In the analysis process the tape was played on an Ampex 601 recorder, the electrical signal was passed through an impedance matching box and into the recorder circuit of a slightly modified "Sona-Graph". Broad band spectrograms were then made of each sentence on the tape, and, for informant #1, narrow band sections were made at target positions of the vowels, i.e. one section if a monothong, two sections if a diphthong, etc. Information recorded from the spectrograms was as follows: (a) Sentence Duration (b) Vowel Segment Duration (c) Duration of Nasal Consonant Following Vowel (d) Frequency values of F_1, F_2, and F_3 for each vowel discernible in the target word.

In the initial stages of the analysis of informant #1's speech, wide band sections were made of the vowels. This procedure was discontinued since the wide band sections did not contribute to the location of formants.

Midway through the processing of the speech from male informant #1, it was observed that the sectioner of the "Sona-Graph" was not giving a display for the temporal point desired but was sectioning early by approximately 10 milliseconds. At the same time it was noted that the left margin of the section display was uneven. These two difficulties were eliminated by strengthening the return spring on the mechanical micro-switch which operates the sectioning gate.

In the analysis of female informant #2's speech, only broad band spectrograms are being used, a decision having been made to no longer take narrow band sections of the vowels. This decision was based on a test comparison between formants determined from broad band only, and formants determined by broad band spectrograms plus narrow band sections. The improvement in accuracy from using narrow band sections in addition to the broad band spectrograms was negligible and within the accuracy limits of the "Sona-Graph" device. The narrow band sections were therefore discontinued to reduce the large amount of time that it takes to process speech by spectrographic analysis. In analysing the high pitched speech of informant #2 it was found that the higher formants were so attenuated that they were no longer evidenced on the spectrograms. As is well known, this problem is quite common with a female voice. It was therefore necessary to build and install a pre-filter which high shaped the electrical signal prior to entering the "Sona-Graph" recorder circuit. It was then possible to obtain clear representations of higher formants without having the low frequency area excessively burned and obscured.

An attempt was also made to obtain reasonably accurate formant values from narrow band spectrograms of female informant #2's speech by modifying the narrow band sampling circuit of the "Sona-Graph" to increase the automatic gain control effect on the output. After a series of adjustments this project was abandoned as impractical.

Results:

As might be noted on Graphs 1 and 2, the mid vowels of Informant #1 tend toward being central rather than markedly front or back. Also, the phone represented as an unrounded central, mid, close, retroflexed vowel might be better represented as an unrounded, back, high, close, retroflexed vowel on the basis of the speech of informant #1.

It might be mentioned that there are considerable differences between the notation used by Hartman and the IPA notation listed below these symbols in the appendices There are instances of calling "different" Hartman symbols the "same" sound, and conversely, calling "same" Hartman symbols "different" sounds. Although the IPA judgements were not made on the basis of the speech of informant #1, but rather on the speech of other native speakers of Mandarin, the physical measurements of informant #1's speech would seem to support these changes.

In addition to the vowel formant values extracted from the speech of informant #1, a preliminary analysis was made on the relation between vowel segment duration and the duration of the following nasal consonant (/n/ or /ŋ/). The results of this preliminary analysis are presented in Appendix #3 but, like all results presented in this paper, must be understood to be highly tenative due to the limited number of samples which were available for analysis and the fact that all data came from a single speaker.

As can be observed in appendix #3 there is no marked effect upon the ratio of nasal ending duration/preceding vowel duration as a function of tone. It therefore appears that although vowel duration is a function of tone, the nasal endings are similarly affected by tone so that the ratio is preserved. One apparent consistency is that the ratio for the velar nasal is always larger than the corresponding ratio for the alveolar nasal, the difference often being approximately a factor of 2. It also appears that the ratio decreases as the amount of physical movement necessary for vowel production increases, as might be expected.

Another preliminary analysis was made to determine the effect on "final" duration due to tone, initial, terminal phoneme, and the number of vowels in the syllable nucleus. Duration of the final (vowel plus terminal nasal, if any) was normalized with respect to sentence duration for all syllables for which all four tones were present and these ratios were investigated. There were a total of 540 syllables (135 for each of the four tones). To investigate the effect of the initial phoneme, these 540 syllables were divided into five groups: those beginning with plosives, sonorants, fricatives, affricates, and vowels, and the average ratio for each group was recorded. To investigate the effect of the terminal phoneme, the samples were divided into three groups: those ending in a vowel, /n/, or /ŋ/, and the average ratio for each group was recorded. To investigate the effect of the number of vowels in the syllable nucleus, the samples were divided into three group monothongs, dipthongs, and triphthongs. The only further division made was on the basis of tone.

The results of this analysis were rather inconclusive but the following generalities appear to be worthy of mention:

(1) The normalized durations of finals for tones 1, 3, and 4 averaged approximately the same.

(2) The normalized durations of finals for tone 2 averaged approximately 10% longer than for the other three tones.

(3) An initial plosive serves to reduce the final duration.

(4) Having no initial serves to increase final duration.

(5) In order of increasing duration of final:

(a) Monothongs (b) Triphthongs (c) Diphthongs

In (5) it might appear that (b) and (c) have a rather unusual order but this can be explained by the fact that triphthongs can never take a nasal ending whereas diphthongs frequently do.

Note:

All data from this experiment as well as notes and figures on the preliminary analyses are kept on file at this office for future comparisons with larger samples of data.

The figure of 1192 words differs from the figure given in previous publications (1178) due to the fact that fourteen words were added after the basic table was compiled.

Appendix #1*

	Formant Frequency		No. of Samples
-ə (ɛ)	F_1	858	39
	F_2	1213	52
	F_3	2338	44
-ja (ʌ)	F_1	876	12
	F_2	1247	15
	F_3	2320	12
-wa (ɑ)	F_1	855	16
	F_2	1222	17
	F_3	2295	13
-aj (a)	F_1	838	47
	F_2	1332	49
	F_3	2309	44
-aw (ɔ)	F_1	783	48
	F_2	1154	61
	F_3	2195	52
-jaw (ʌ)	F_1	642	33
	F_2	1174	35
	F_3	2196	26
-waj (a)	F_1	728	14
	F_2	1278	16
	F_3	2252	13
-an (ə)	F_1	858	59
	F_2	1327	63
	F_3	2340	54
-ə (ɑ)	F_1	801	47
	F_2	1235	56
	F_3	2341	51
-waŋ (ʌ)	F_1	685	15
	F_2	1114	20
	F_3	2045	11
-jaŋ (a)	F_1	743	18
	F_2	1329	20
	F_3	2335	15
-jin (ɛ)	F_1	446	32
	F_2	1649	32
	F_3	2474	29
-yan (ɛ)	F_1	481	13
	F_2	1474	15
	F_3	2331	7
-wan (a)	F_1	764	23
	F_2	1320	36
	F_3	2333	23
-ë (ɘ)	F_1	467	34
	F_2	1199	34
	F_3	2165	32
-jE (ɛ)	F_1	462	31
	F_2	1653	29
	F_3	2552	22
-ʌ (ʌ)	F_1	468	55
	F_2	1099	55
	F_3	(2219) 2340	(41) 33

*Data from Male Informant #1.

	Formant Frequency		No. of Samples
-yE	F_1	446	14
(ɛ)	F_2	1557	13
	F_3	2357	9
-ɐj	F_1	484	23
(ɐ)	F_2	1652	17
	F_3	2492	18
-ɐw	F_1	465	47
(o)	F_2	1043	45
		1024	42
	F_3	2218	22
-iɐw	F_1	330	1
(ʊ)	F_2	1000	1
	F_3	(1600)	1
-wɐj	F_1	437	10
(e)	F_2	1872	9
	F_3	2621	7
-wIj	F_1	396	8
(e)	F_2	1911	7
	F_3	---	--
-ɐn	F_1	443	36
(ə)	F_2	1396	38
	F_3	2196	27
wɐn	F_1	336	3
(ə)	F_2	1400	2
	F_3	2025	2
-wën	F_1	488	4
(ə)	F_2	1268	4
	F_3	2083	3
-ɐr	F_1	448	43
(ə)	F_2	1199	44
	F_3	2195	41
		(2726)	(8)
		2085	32
wɐj	F_1	475	2
(ə)	F_2	1265	2
	F_3	2300	1
-z˙	F_1	326	10
(+)	F_2	1132	10
	F_3	2750	1
-ɝ˙	F_1	332	13
(ɝ)	F_2	972	10
	F_3	2103	6
-i	F_1	277	42
(i)	F_2	2015	40
	F_3	3128	33
-u	F_1	333	66
(u)	F_2	937	62
	F_3	1604	23
-y	F_1	269	20
(y)	F_2	1907	20
	F_3	2354	14
-in	F_1	294	25
(i)	F_2	(921)	(21)
		1977	24
	F_3	2777	18
-yn	F_1	291	12
(Y)	F_2	(920)	(8)
		1851	11
	F_3	2425	6

	Formant Frequency		No. of Samples
-[illegible]	F_1	304	37
	F_2	(930)	30
(i)		1980	31
	F_3	2799	29
-[illegible]	F_1	397	38
(a)	F_2	1077	37
	F_3	1916	16
-[illegible]	I_1	316	11
	F_2	948	11
(u)	F_3	1923	7

Appendix #2**

	Formant Frequency		No. of Samples
-å (ɑ)	F_1	1319	15
	F_2	1684	1[illegible]
	F_3	2746	[illegible]
-u (u)	F_1	417	14
	F_2	950	11
	F_3	3500	2
-ɚ (ə)	F_1	975	13
	F_2	1485	13
	F_3	2191	7
-a͡ɪ (ʌ)	F_1	1242	13
	F_2	1673	12
	F_3	2755	10
-ən (ə)	F_1	915	13
	F_2	2072	13
	F_3	3076	13
-an (a)	F_1	1122	13
	F_2	1771	13
	F_3	2822	12
-o͡w (o)	F_1	450	3
	F_2	1030	3
	F_3	3116	3
-o͡w (ʌ)	F_1	1056	9
	F_2	1434	8
	F_3	2860	4
-ej (ɛ)	F_1	634	7
	F_2	2255	9
	F_3	3105	10
-aj (a)	F_1	1200	9
	F_2	1763	10
	F_3	2602	9
-wɔ (ɔ)	F_1	1006	8
	F_2	1900	2
	F_3	3125	2

**Data from Female Informant #2

Appendix #3

Report on ratio of nasalized ending duration/preceding vowel duration

The data for the results listed below were obtained from the spectrograms of the speech of informant #1. The word "ratio" will refer to the ratio of nasalized ending duration/preceding vowel duration. All grouping of vowels by vowel position is based on Hartman's notation. Also in this appendix, the word "monothong" will refer to a single vowel such as /i/ or /a/ whereas the word "diphthong" will refer to a combination of vowel plus semi-vowel such as /ja/ or /aj/.

I.

Nasal	Tone	# of Samples	Average Ratio	
n	1	28	.8666	
n	2	28	.8061	(Data taken only from
n	3	28	.7979	syllables in which all
n	4	28	.8109	four tones were present
ŋ	1	21	1.4082	
ŋ	2	21	1.4309	
ŋ	3	21	1.4820	
ŋ	4	21	1.2551	

II. Overall ratios compared:

Nasal	# of Samples	Average Ratio
n	270	.8845
ŋ	230	1.7266

III. Average ratios as a function of vowel position for monothongs:

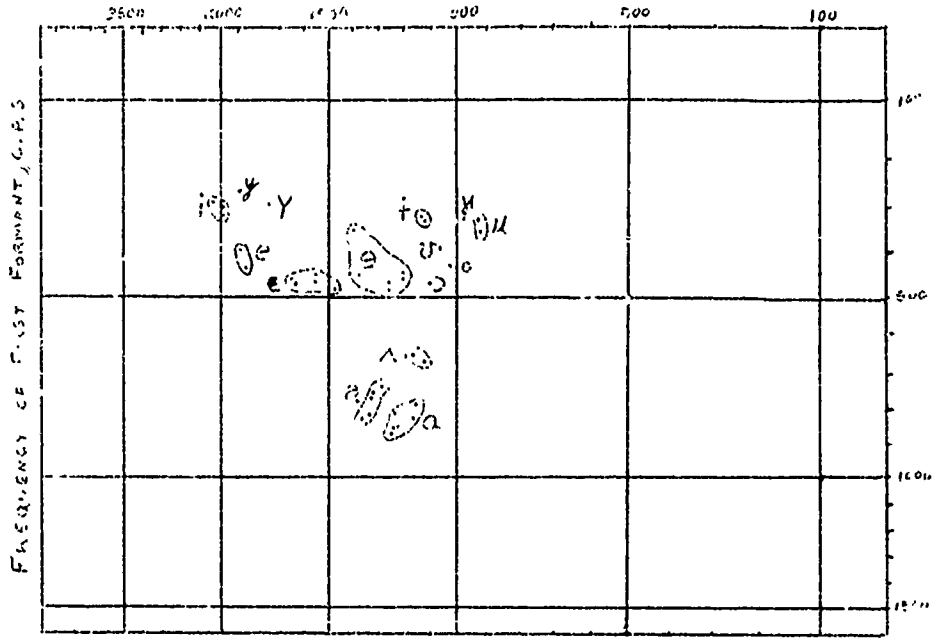
Vowel position	Nasal	# of Samples	Ratio
High	n	55	1.2228
High	ŋ	57	2.1655
Mid	n	49	1.1793
Mid	ŋ	56	2.6528
Low	n	65	.7184
Low	ŋ	35	.9326

IV: Average ratios as a function of vowel position for dipthongs:

Vowel position	Nasal	# of Samples	Ratio
High to Low	n	70	.6070
	ŋ	40	.8394
High to Mid	n	38	.7083
	ŋ	15	.8946
Front to Front	n	24	.5677
	ŋ	6	.8277
Back to Back	n	50	.6005
	ŋ	22	.8618
Back to Front	n	14	.9493
	ŋ	2	1.0955
Front to Back	n	20	.6232
	ŋ	25	1.1377

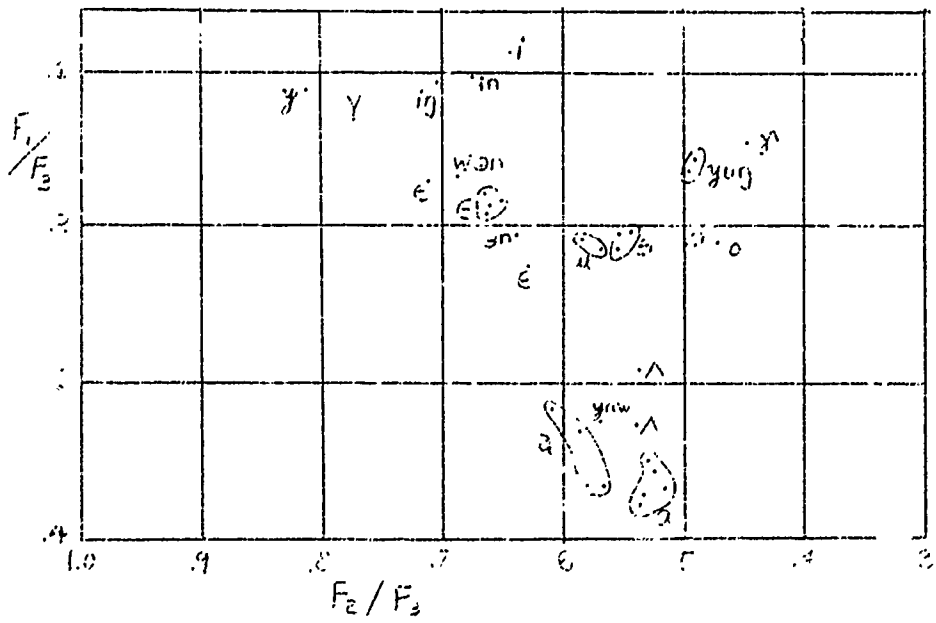
All Diphthongs

Nasal	# of Samples	Ratio
n	108	.6426
ŋ	54	.9920



Mel scale Phonemes are in IPA notation Data from Informant #1

Graph #1



Finals are in IPA notation Data from Informant #1

Graph #2

TONE PERCEPTION EXPERIMENT

K. P. Li

Introduction:

In Mandarin Chinese, whenever a tone 3 is followed by another tone 3, it becomes a modified tone 3 whose quality has long been a subject of debate. To most people, this modified tone 3 is not clearly distinguishable from tone 2. However, some linguists have suggested that its pitch may be on a lower level than that of tone 2.[1] This experiment has been designed to determine whether, in fact, native speakers are able to distinguish between this modified tone 3 and tone 2 in a controlled environment..

Procedures:

A set of disyllabic pairs, with varying degrees of meaningfulness, were chosen. The forms in each pair were phonemically identical, differing only in the tone of the first syllable: tone 2 in the first of each pair, tone 3 in the second; the second syllable was always tone 3, e.g., mái-mǎ (埋馬 bury a horse) mǎi-mǎ (買馬 buy a horse); yú-fǎ (漁法 law of fishing) yǔ-fǎ (語法 syntax). 130 such pairs were collected (see list in Table IV).

Two Mandarin speakers (No. 2 and No. 3 in table on p. 8) were chosen to make recordings[2] of a list in which each disyllable of each pair appeared twice in random order, yielding 520 items read. Corresponding to each item was a pause of about 3 seconds during which the listener was to make his decision and circle the appropriate disyllable. Both the reading list and answer sheet were written in Chinese characters.[3]

Before administering the actual test, a sample test of 10 similar

---

[1] See Hockett, C. F., "Peiping Morphophonemics," Language 26 (1950), pp. 63-85, and Martin, S. E., "Problems of Hierarchy and Indeterminacy in Mandarin Phonology", Bulletin of the Institute of History and Philology, Academia Sinica, No. 29, Part I, pp. 209-230.

[2] The acoustic conditions for recording were the same as those shown on p. 8.

[3] An answer sheet in the Pin-Yin Romanization system has been prepared, but we have not as yet had any subject use it.

Table I

Listener	No. of Correct Responses for Tone 2	No. of Correct Responses for Tone 3	Total No. of Correct Responses	Percentage of Correct Responses
1.	117	144	261	50.2
2.	128	135 (1)*	263 (1)*	50.7
3.	149	110	259. .	49.8
4.	137	138	275	52.9
5.	161	112	273	52.5
6.	184 (4)	93 (6)	277 (10)	54.3
7.	132	135	267	51.3
8.	107 (50)	108 (57)	215 (107)	52.0
9.	172	113	285	54.8
10.	126	130	256	49.2
11.	171	105	275	52.9
12.	135	134	269	51.7
13.	141	126	267	51.3
14.	144	138	282	54.2
Total	2004 (54)	1721 (6.)	3725 (118)	52.0
Speaker No. 2	209	87 (4)	296 (4)	56.9%
Speaker No. 3	131	209	340	67.3%

*The number in parenthesis represents the number of undecided.

Table II

Final of First Syllable	No. of Pairs	Tone 2 No. of Correct Responses	Tone 2 No. of Undecided	Tone 3 No. of Correct Responses	Tone 3 No. of Undecided	Total Percentage of Correct Responses
-iu, -ou	6	122	3	61	2	55.3
-iou	7	136	4	100	3	61.2
-uo, -o	9	164	3	93	4	51.6
-ai	5	76	4	53	2	47.1
-ao, -iao	8	103	6	128	3	52.5
-i	17	210	6	280	11	52.4
-u	14	227	4	165	9	50.9
-ei	3	34	1	52	1	51.7
-ü	4	46	0	72	4	53.7
-a, -ia	8	98	4	118	3	48.9
-üe, -ie	3	24	0	62	1	51.5
-er	1	7	1	19	0	67.7
-z	1	23	1	6	0	52.8
-ian, üan	10	144	3	140	5	51.4
-an, -wan	10	162	[illegible]	119	7	49.5
-iang,-ang -wang	11	189	5	119	6	50.9
-ing, üng	5	89	2	53	2	51.5
-en, -eng	3	56	2	31	0	52.4
-in	2	25	1	29	0	48.6
-un, -ung -ün	3	69	1	21	1	54.2
	130	2004	54	1721	64	52.0

disyllables, with either different tones or different consonants (also listed in Table IV) were arranged for testing the recording conditions and the dialect background of the listeners.

Participants:

14 Chinese Mandarin instructors (12 men and 2 women) of the Institute of Far Eastern Languages at Yale University were chosen as listeners.[4] The two speakers also took the test, each listening to his own tape.

Results:

Only three of the 14 listeners had one error each in the sample test (an error percentage of 2.1%); the others had none.

From the test itself, the total percentage of correct responses for distinguishing modified tone 3 and tone 2 was only 52.0%. The number of correct responses and the percentages for individual listeners and for the different finals of the first syllable are shown in Tables I and II respectively. The results of the speakers themselves are better (56.9% for No. 2 speaker and 67.3% for No. 3 speaker). However, even these percentages are low from a statistical point of view.

All disyllable pairs were then classified into three groups (shown in Table IV). Group A contained 45 pairs. In each pair, the members were judged to have an equal degree of meaningfulness and of frequency of occurrence, e.g., fén-chǎng (墳場 grave yard) and fěn-chǎng (粉廠 flour factory); níu-yǐng (牛影 cow's shadow) and nǐu-yǐng (扭影 twist shadow). Group B consisted of 40 pairs in which the disyllable with tone 2 had a higher degree of frequency of occurrence than that with tone 3, e.g., cáo-fǔ (曹府 hell) and cǎo-fǔ (草腐 grass rotten); zhí-jiǎo (直角 right angle) and zhì-jiǎo (紙角 paper corner). Group C contained 45 pairs in which the disyllable with tone 3 was of more common occurrence, e.g., bí-měi (鼻美 nose beauty) and bǐ-měi (比美 beauty competition); yóu-hǎo (油好 oil good) and yǒu-hǎo (友好 friendship). The percentages of correct responses for each group is shown in the following table:

---

[4] We wish to express our appreciation to the Institute of Far Eastern Languages of Yale University for their help.

Table III

	correct response as tone 2 read in %	correct response as tone 3 read in %	total correct response in %
Group A	57.9	46.6	52.3
Group B	82.0	25.7	53.8
Group C	30.7	69.5	50.1
Total	55.9	48.1	52.0

Conclusion:

From the results of the experiement, it is evident that Mandarin speakers are not able to distinguish a modified tone 3 from a tone 2, even in the case of speakers listening to their own recordings.

Furthermore, from the results of the grouping in the above table, it can be seen that the decision of the listeners is affected by the meaningfulness of the disyllable. They have a tendency to choose the one which has more meaning or more frequent usage, in spite of the fact that they are instructed in the beginning of the test to disregard meaning and to base their decision solely on pronunciation.

Suggestion:

We are planning to use the pitch extractor of a vocoder to measure the pitch of these disyllables. At present we are improving the performance of the pitch extractor to satisfy the specification for Mandarin which has a fast rate of change of pitch level and intensity. As soon as it becomes feasible, we may have the result from the mechanical measurement to subjoin the argument of the difference of modified tone 3 and tone 2.[5]

---

[5] cf. a related experiment done by Yao Shen, Jessica C.Y. Chao, and Giles Peterson of the University of Michigan in "Some Spectrographic Light on Mandarin Tone-2 and Tone-3", Study of Sounds, IX, Phonetic Society of Japan, 1961.

Table IV

I. 10 Sample Disyllabic Pairs

肥灰	fēi-huī	灰飛	huī-fēi
勸世	juàn-shì	勸世	quàn-shì
屠殺	tú-shā	毒殺	dú-shā
至尊	zhì-zūn	自尊	zì-zūn
談情	tán-qíng	彈琴	tán-qín
早上	zhǎo-shàng	草上	chǎo-shàng
使用	shǐ-yòng	實用	shí-yòng
吸食	xī-shí	惜食	xí-shí
搬演	bān-yǎn	板演	bǎn-yǎn
師娘	shī-niáng	石娘	shí-niáng

II. 130 Tested Disyllabic Pairs

ba-shǒu	板手	把手	C	du-shǒu	毒手	賭手	B
bai-shǒu	白首	百首	B	du-yǒu	獨有	賭友	A
bai-shǒu	白手	擺手	A	duo-zǒu	拿走	躲走	A
bao-shǒu	薄手	保守	C	er-yǔ	兒語	耳語	A
bi-měi	鼻美	比美	C	fa-kuǎn	罰款	法款	A
cai-cǎo	材草	採草	A	fan-guǐ	鬼	反鬼	A
cai-nǚ	才女	彩女	L	fang-shǒu	防守	紡手	B
cao-fǔ	曹府	草腐	B	fen-chǎng	墳場	粉廠	A
chan-cǎo	纏草	產草	C	gu-jiǔ	骨酒	古酒	A
chang-jiǔ	長久	場酒	B	guo-chǎn	國產	果產	A
chang-yuǎn	長遠	場遠	B	guo-chǐ	國恥	果尺	B
chou-zhě	仇者	醜者	A	han-kǔ	寒苦	喊苦	A
chun-yǒu	純友	蠢友	A	han-shǔ	寒暑	喊鼠	B
da-guǐ	答鬼	打鬼	C	han-yǎng	寒癢	喊癢	B
da-zuǐ	答嘴	打嘴	A	han-yǔ	寒雨	喊雨	B
dǎ-huǐ	敲毀	誑燬	C	hao-mǎ	豪馬	好馬	C

hao-yǐn	豪飲	好飲	B	qi-cǎo	奇草	起草	C
hu-hǒu	胡吼	虎吼	C	qi-gǎo	奇稿	起稿	C
hu-yǎn	糊眼	虎眼	A	qi-kěn	其肯	豈肯	C
huo-kǒu	活口	火口	C	qi-mǎ	騎馬	起碼	A
huo-shuǐ	活水	火水	A	qi-sǐ	騎死	起死	C
huo-yǎn	活眼	火眼	A	qin-shì	琴室	寢室	A
huo-yǒu	活友	夥友	C	qing-wǎng	情網	請往	B
ji-sǐ	急死	擠死	A	qing-yǔ	晴雨	請雨	B
jia-lǐ	夾裡	假理	B	qu-shuǐ	渠水	取水	A
lao-guǐ	撈鬼	老鬼	C	ran-qǐ	燃起	染起	B
leng-jiǎo	稜角	冷腳	B	ren-sǐ	人死	忍死	B
li-fǎ	釐法	理髮	C	shou-zhǔ	熟煮	守主	C
li-jiě	離解	理解	C	tao-hǎo	逃好	討好	C
lian-pǔ	連譜	臉譜	C	ti-tǒng	提桶	體統	C
liang-jiǎo	量角	兩角	A	tong-bǎn	銅板	桶板	A
liao-jiě	遼姐	了解	C	tu-gǎi	塗改	土改	A
ling-bǎo	靈寶	領寶	A	wa-lěng	娃冷	瓦冷	A
ling-jiǎo	菱角	領角	B	wan-hǎo	玩好	碗好	A
liu-wǔ	劉五	柳五	A	wan-jǐng	玩景	晚景	C
liu-yǐng	留影	柳影	B	wan-qǐ	玩起	晚起	C
ma-shǎo	麻少	馬少	A	wang-sǐ	王死	枉死	A
mai-mǎ	埋馬	買馬	A	wu-bǐ	無比	五筆	B
mei-jiǔ	沒酒	美酒	C	wu-cǎi	無彩	五彩	A
mei-nǚ	煤女	美女	C	wu-kǒu	無口	五口	C
mei-yǎn	眉眼	美眼	B	wu-lǐ	無理	五里	A
mi-jiǔ	迷酒	米酒	A	wu-nǚ	吳女	舞女	C
mian-sǐ	棉死	免死	C	xi-jiǔ	席酒	喜酒	C
mo-guǐ	魔鬼	抹鬼	B	xiang-xiě	詳寫	想寫	C
mo-kǒu	磨口	抹口	A	xie-gǎo	鞋稿	寫稿	C
ni-gǔ	泥鼓	擬古	C	xue-chǐ	學尺	雪恥	C
ni-hǎo	泥好	你好	C	ya-pǔ	牙譜	雅譜	A
nian-zhǐ	年紙	撚指	A	yan-kǒu	鹽口	掩口	C
niu-jiǎo	牛角	扭腳	B	yan-lǐ	鹽里	眼裡	A
niu-yǐng	牛影	扭影	A	yan-shǔ	鹽署	偃鼠	B

yan-shuǐ	鹽水	眼水	B
yan-yǔ	言語	眼語	B
yang-hǔ	楊虎	養虎	A
yang-jiǎo	羊角	仰角	B
yang-nǚ	洋女	養女	C
yang-shǒu	揚首	仰首	C
yang-zhǐ	洋紙	仰止	B
ye-hǎo	爺好	也好	C
yi-biǎo	儀表	乙表	B
yi-lǎo	遺老	倚老	A
yin-yǔ	淫雨	隱語	A
yong-bǎo	庸保	永保	C
you-bǐng	油餅	有餅	B
you-guǐ	油鬼	有鬼	C
you-hǎo	油好	友好	C
you-jǐng	油井	有井	B
you-shǒu	游手	有手	B
you-wǒ	由我	有我	A
you-yǒng	游泳	有勇	B
yu-fǎ	漁法	語法	C
yu-gáng	漁港	雨港	A
yu-shuǐ	魚水	雨水	C
yuan-zhǔ	園主	遠主	B
yuan-zǔ	源阻	遠阻	C
yun-duǒ	雲朵	允躲	B
zhao-huǒ	着火	找火	B
zhi-jiǎo	直角	紙角	B
zhu-mǐ	竹米	煮米	A
zhu-sǔn	竹筍	煮筍	B
zhu-wǎn	竹碗	煮碗	B

	ㄚ	ㄛ	ㄞ	ㄟ	ㄠ	ㄡ	ㄢ	ㄣ	ㄤ	ㄥ	ㄨ	ㄧ	ㄧㄚ	ㄧㄝ	ㄧㄠ	ㄧㄡ	ㄧㄢ	ㄧㄣ	ㄧㄤ	ㄧㄥ
	a	o	ai	ei	au	ou	an	en	ang	ong	u	i	ia	ie	iau	iu	ian	in	iang	ing
ㄅ b	疤拔把霸	玻勃跛簸	掰白百敗	悲北倍	包雹保報		班板半	奔本笨	幫綁棒	崩蹦	醭補布	逼鼻比閉		鱉別癟憋	標表		邊扁變	賓稟鬢		兵餅病
ㄆ p	趴爬怕	坡婆叵破	拍排派	披陪配	抛袍跑砲	剖	潘盤判	噴盆坌	乓旁嗙胖	烹棚捧碰	撲葡普舖	批皮匹僻		撇苤	飄瓢漂票		偏便騙	拼貧品聘		乒平聘
ㄇ m	媽麻馬罵	摸魔抹沫	埋買賣	眉美妹	貓毛卯帽	謀某貿	顢瞞滿慢	悶門燜	忙蟒	盟猛孟	摹母木	瞇迷米蜜		咩滅	喵苗秒廟	謬	棉免面	民敏		明皿命
ㄈ f	發罰法琺	佛		非肥匪費		蜉否	翻煩反飯	分墳粉奮	方房訪放	風馮諷鳳	夫扶府父									

	ㄚ	ㄜ	ㄞ	ㄟ	ㄠ	ㄡ	ㄢ	ㄣ	ㄤ	ㄥ	ㄨ	ㄨㄚ	ㄨㄛ	ㄨㄞ	ㄨㄟ	ㄨㄢ	ㄨㄣ	ㄨㄤ	ㄨㄥ
	a	e	ai	ei	ao	ou	an	en	ang	eng	u	ua	uo	uai	ui	uan	un	uang	ong
ㄉ d	搭答打大	德	獃歹代	得	刀倒島到	兜斗豆	單膽但		當黨蕩	登等凳	督獨賭度		多奪朵惰		堆對	端短段	敦盹頓		冬懂動
ㄊ t	他塔踏	特	胎枱泰		叨桃討套	偷投透	貪痰毯炭		湯唐躺燙	騰	禿圖土兔		拖駝妥唾		推頹腿退	湍團	吞屯褪		通同統痛
ㄋ n	哪拿那納	訥	奶耐	餒內	撓腦鬧	耨	男赧難	嫩	囊攮齉	能	奴努怒		挪娜諾			暖	嫩		農弄
ㄌ l	拉剌喇辣	勒肋	來賴	勒雷壘類	撈牢老烙	摟樓簍漏	藍懶爛		狼朗浪	稜冷愣	嚕爐魯鹿		囉螺裸駱			灤卵亂	掄輪論		龍攏弄

	ㄧ	ㄧㄚ	ㄧㄝ	ㄧㄠ	ㄧㄡ	ㄧㄢ	ㄧㄣ	ㄧㄤ	ㄧㄥ	ㄩ	ㄩㄝ	ㄩㄢ	ㄩㄣ	ㄩㄥ
	i	ia	ie	iau	iu	ian	in	iang	ing	ü	üe	üan	ün	iung
ㄉ d	低敵底弟		爹疊	雕 掉	丟	顛 點電			丁 頂定					
ㄊ t	梯提體替		貼 鐵帖	挑條窕跳		天田舔			聽停挺					
ㄋ n	泥你膩		捏 聶	鳥尿	妞牛紐謬	蔫年碾念	您	娘 釀	寧擰濘	女	虐			
ㄌ l	離里利	倆	咧列	撩遼了料	溜流柳遛	連臉練	鄰凜吝	良兩亮	拎靈領令	驢旅慮	略	孿戀	淋	

	ㄧ	ㄧㄚ	ㄧㄝ	ㄧㄠ	ㄧㄡ	ㄧㄢ	ㄧㄣ	ㄧㄤ	ㄧㄥ	ㄩ	ㄩㄝ	ㄩㄢ	ㄩㄣ	ㄩㄥ
	i	ia	ie	iau	iu	ian	in	iang	ing	ü	üe	üan	ün	iong
ㄉ d	低敵底弟		爹疊	雕 掉	丟	顛 點電			丁 頂定					
ㄊ t	梯提體替		貼 鐵帖	挑條窕跳		天田舔			聽停挺					
ㄋ n	泥你膩		捏 聶	鳥尿	妞牛紐謬	蔫年碾念	您	娘 釀	寧擰濘	女	虐			
ㄌ l	離里利	倆	咧列	撩遼了料	溜流柳遛	連臉練	鄰凜吝	良兩亮	拎靈領令	驢旅慮	略	孿巒	淋	

	ㄚ	ㄜ	ㄞ	ㄟ	ㄠ	ㄡ	ㄢ	ㄣ	ㄤ	ㄥ	ㄨ	ㄨㄚ	ㄨㄛ	ㄨㄞ	ㄨㄟ	ㄨㄢ	ㄨㄣ	ㄨㄤ	ㄨㄥ
	a	o	ai	ei	ao	ou	an	en	ang	eng	u	ua	uo	uai	ui	uan	un	uang	ong
ㄍ g	嘎	哥格葛個	該改蓋	給	高稿告	溝狗够	干趕幹	根哏艮亘	剛港槓	庚耿更	姑骨古故	瓜寡掛	鍋國果過	乖拐怪	歸鬼貴	官管慣	滾棍	光廣逛	功鞏供
ㄎ k	咖卡	科咳可課	開凱慨		尻考靠	摳口扣	刊砍看	肯	康扛抗	坑肯	哭苦褲	誇垮跨	闊	蒯快	虧葵傀愧	寬款	昆捆困	筐狂況	空恐控
ㄏ h	哈蛤	喝河赫	孩海害	黑	蒿豪好號	齁猴吼後	酣寒喊汗	痕狠恨	夯航	亨恆橫	乎胡虎戶	花華化	豁活火禍	懷壞	灰回毀會	歡桓緩換	婚魂混諢	荒黃謊恍	轟紅哄鬨

		ㄓ	ㄔ	ㄕ	ㄖ	ㄗ	ㄘ	ㄙ	ㄚ	ㄛ	ㄜ	ㄝ	ㄞ	ㄟ	ㄠ	ㄡ	ㄢ	ㄣ	ㄤ	ㄥ	ㄦ	ㄧ	ㄨ	ㄩ
		i							a	o	e	ê	ai	ei	ao	ou	an	en	ang	eng	er	yi	wu	yu
ㄓ	zh	知直只志	吃遲尺翅	師示史市	日																			
ㄔ	ch								插閘眨擦	叉茶蹅詫	沙 傻然													
ㄕ	sh								遮摺者澌	車 扯澈	餘蛇捨射		惹熱											
ㄖ	r				日																			
ㄗ	z					齋宅窄債	拆柴	節 色晒																
ㄘ	c																							
ㄙ	s																							
ㄚ	a																							
ㄛ	o																							
ㄜ	e																							
ㄝ	ê																							
ㄞ	ai																							
ㄟ	ei																							
ㄠ	ao																							
ㄡ	ou																							
ㄢ	an																							
ㄣ	en																							
ㄤ	ang																							
ㄥ	eng																							
ㄦ	er																							
ㄧ	yi																							
ㄨ	wu																							
ㄩ	yu																							

		ㄚ	ㄜ	ㄞ	ㄟ	ㄠ	ㄡ	ㄢ	ㄣ	ㄤ	ㄥ	ㄨ	ㄨㄚ	ㄨㄛ	ㄨㄞ	ㄨㄟ	ㄨㄢ	ㄨㄣ	ㄨㄤ	ㄨㄥ
	i	a	e	ai	ei	ao	ou	an	en	ang	eng	u	ua	uo	uai	ui	uan	un	uang	ong
ㄗ z	資 紫 自	紮 雜 軋	責 仄	災 宰 在	賊	糟 鑿 早 造	諏 走 奏	簪 咱 攢 贊	怎	髒 葬	增 贈	租 足 組		嘬 昨 左 坐		嘴 罪	鑽 纂 攥	尊		宗 總 縱
ㄘ c	疵 辭 此 次	擦	策	猜 才 採 菜		操 曹 草	湊	餐 殘 慘 燦	參	蒼 藏	層 蹭	粗 醋		搓 錯		催 脆	躥 竄	村 存 忖 寸		聰 從
ㄙ s	思 死 四	撒 灑 颯	澀	腮 賽		騷 嫂 臊	搜 叟 嗽	三 傘 散	森 滲	桑 嗓 喪	僧	蘇 俗 素		縮 所		雖 隨 碎	酸 算	孫 損 遜		鬆 聳 宋
		a	e	ai	ei	ao	ou	an	en	ang	eng	wu	wa	wo	wai	wei	wan	wen	wang	weng
		啊 阿	額 噁 扼	哀 捱 矮 愛	欸	鏖 熬 襖 奧	毆 偶 漚	安 黯 岸	恩	骯 昂 盎	嗯	屋 吳 五 務	挖 娃 瓦 襪	窩 我 臥	歪 外	威 圍 偉 畏	彎 完 晚 萬	溫 文 穩 問	汪 王 網 忘	翁 甕

ㄦ	er	而耳二

ㄧㄞ	yai	崖

	ㄧ	ㄧㄚ	ㄧㄝ	ㄧㄠ	ㄧㄡ	ㄧㄢ	ㄧㄣ	ㄧㄤ	ㄧㄥ	ㄩ	ㄩㄝ	ㄩㄢ	ㄩㄣ	ㄩㄥ
	i	ia	ie	iao	iu	ian	in	iang	ing	u	ue	uan	un	iong
ㄐ j	基急己記	加莢甲架	街傑解界	交嚼腳叫	究久舊	奸簡件	斤緊近	江講匠	京井竟	居局舉句	噘決 [illegible]	捐捲倦	軍窘菌	扃炯郡
ㄑ q	妻其起器	掐卡恰	切茄且竊	敲橋巧竅	秋球	千前淺欠	親琴寢沁	槍牆搶嗆	輕晴請慶	區渠娶去	缺瘸闕	圈權犬勸	羣雋	穹窮
ㄒ x	希習喜戲	蝦霞下	歇鞋寫謝	消肴曉笑	休朽秀	先賢顯現	新信	香祥想向	星形醒幸	須徐許緒	靴學雪穴	宣懸選漩	勳巡訓	兄雄
	yi	ya	ye	yao	yo	yan	yin	yang	ying	yu	yue	yu...	yun	yong
	衣疑以易	丫牙雅訝	椰爺也夜	妖搖咬要	憂由友又	煙延演厭	音銀飲印	央羊養樣	英營影硬	迂餘語預	約越	冤園遠院	暈雲允韻	雍庸永用

INVENTORY OF MANDARIN SYLLABLES ACCORDING TO THE PINYIN ALPHABET

INITIAL \ FINAL	Syllable Inventory According to the PINYIN Alphabet - 1											
	ER	A	E	AI	EI	AO	OU	AN	EN	ANG	ENG	ONG
	er	a	e	ai	ei	ao	ou	an	en	ang	eng	
B		ba		bai	bei	bao		ban	ben	bang	beng	
P		pa		pai	pei	pao	pou	pan	pen	pang	peng	
M		ma		mai	mei	mao	mou	man	men	mang	meng	
F		fa			fei		fou	fan	fen	fang	feng	
D		da	de	dai	dei	dao	dou	dan		dang	deng	dong
T		ta	te	tai		tao	tou	tan		tang	teng	tong
N		na	ne	nai	nei	nao	nou	nan	nen	nang	neng	nong
L		la	le	lai	lei	lao	lou	lan		lang	leng	long
Z		za	ze	zai	zei	zao	zou	zan	zen	zang	zeng	zong
C		ca	ce	cai		cao	cou	can	cen	cang	ceng	cong
S		sa	se	sai		sao	sou	san	sen	sang	seng	song
ZH		zha	zhe	zhai	zhei	zhao	zhou	zhan	zhen	zhang	zheng	zhong
CH		cha	che	chai		chao	chou	chan	chen	chang	cheng	chong
SH		sha	she	shai	shei	shao	shou	shan	shen	shang	sheng	
R			re			rao	rou	ran	ren	rang	reng	rong
G		ga	ge	gai	gei	gao	gou	gan	gen	gang	geng	gong
K		ka	ke	kai		kao	kou	kan	ken	kang	keng	kong
H		ha	he	hai	hei	hao	hou	han	hen	hang	heng	hong

Syllable Inventory according to the PINYIN Alphabet - 2

INITIAL \ FINAL	I	IA	IE	IAO	IOU	IN	ING	IAN	IANG	IONG
	yi	ya	ye	yao	you	yin	ying	yan	yang	yong
B	bi		bie	biao		bin	bing	bian		
P	pi		pie	piao		pin	ping	pian		
M	mi		mie	miao	miu	min	ming	mian		
F										
D	di		die	diao	diu		ding	dian		
T	ti		tie	tiao			ting	tian		
N	ni		nie	niao	niu	nin	ning	nian	niang	
L	li	lia	lie	liao	liu	lin	ling	lian	liang	
Z	zi									
C	ci									
S	si									
ZH	zhi									
CH	chi									
SH	shi									
R	ri									
G	ji	jia	jie	jiao	jiu	jin	jing	jian	jiang	jiong
K	qi	qia	qie	qiao	qiu	qin	qing	qian	qiang	qiong
H	xi	xia	xie	xiao	xiu	xin	xing	xian	xiang	xiong

INITIAL \ FINAL 35 / 18	Syllable Inventory According to the PINYIN Alphabet - 3													
	U	UO	UA	UAI	UEI	UAN	UN	UANG	UENG	Ü	ÜE	ÜAN	ÜN	405
	wu	wo	wa	wai	wei	wan	wen	wang	weng	yu	yue	yuan	yun	34
B	bu	bo												16
P	pu	po												17
M	mu	mo												18
F	fu	fo												9
D	du	duo			dui	duan	dun							21
T	tu	tuo			tui	tuan	tun							19
N	nu	nuo				nuan	nun			nü	nüe			25
L	lu	luo				luan	lun			lü	lüe	lüan	lün	27
Z	zu	zuo			zui	zuan	zun							17
C	cu	cuo			cui	cuan	cun							16
S	su	suo			sui	suan	sun							16
ZH	zhu	zhuo	zhua	zhuai	zhui	zhuan	zhun	zhuang						20
CH	chu	chuo	chua	chuai	chui	chuan	chun	chuang						19
SH	shu	shuo	shua	shuai	shui	shuan	shun	shuang						19
R	ru	ruo			rui	ruan	run							14
G	gu	guo	gua	guai	gui	guan	gun	guang		ju	jue	juan	jun	33
K	ku	kuo	kua	kuai	kui	kuan	kun	kuang		qu	que	quan	qun	32
H	hu	huo	hua	huai	hui	huan	hun	huang		xu	xue	xuan	xun	33

COMPARATIVE TABLE OF PINYIN, YALE
WADE-GILES, ZHUYIN ZIMU
AND GWOYEU ROMATZYH (TONAL SPELLING) SYSTEMS*

PY	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
a	a	a	ㄚ	a	ar	aa	ah
ai	ai	ai	ㄞ	ai	air	ae	ay
an	an	an	ㄢ	an	arn	aan	ann
ang	ang	ang	ㄤ	ang	arng	aang	anq
ao	au	ao	ㄠ	au	aur	ao	aw
ba	ba	pa	ㄅㄚ	ba	bar	baa	bah
bai	bai	pai	ㄅㄞ	bai	bair	bae	bay
ban	ban	pan	ㄅㄢ	ban	barn	baan	bann
bang	bang	pang	ㄅㄤ	bang	barng	baang	banq
bao	bau	pao	ㄅㄠ	bau	baur	bao	baw
bei	bei	pei	ㄅㄟ	bei	beir	beei	bey
ben	ben	pen	ㄅㄣ	ben	bern	been	benn
beng	beng	peng	ㄅㄥ	beng	berng	beeng	benq
bi	bi	pi	ㄅㄧ	bi	byi	bii	bih
bian	byan	pien	ㄅㄧㄢ	bian	byan	bean	biann
biao	byau	piao	ㄅㄧㄠ	biau	byau	beau	biaw
bie	bye	pieh	ㄅㄧㄝ	bie	bye	biee	bieh
bin	bin	pin	ㄅㄧㄣ	bin	byn	biin	binn
bing	bing	ping	ㄅㄧㄥ	bing	byng	biing	binq
bo	bwo	po	ㄅㄛ	bo	bor	boo	boh
bu	bu	pu	ㄅㄨ	bu	bwu	buu	buh
ca	tsa	ts'a	ㄘㄚ	tsa	tsar	tsaa	tsah
cai	tsai	ts'ai	ㄘㄞ	tsai	tsair	tsae	tsay
can	tsan	ts'an	ㄘㄢ	tsan	tsarn	tsaan	tsann
cang	tsang	ts'ang	ㄘㄤ	tsang	tsarng	tsaang	tsanq
cao	tsau	ts'ao	ㄘㄠ	tsau	tsaur	tsao	tsaw

*See p. 63 for references for these systems.

??	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
ce	tse	ts'e	ㄘㄜ	tse	tser	tsee	tseh
cen	tsen	ts'en	ㄘㄣ	tsen	tsern	tseen	tsenn
ceng	tseng	ts'eng	ㄘㄥ	tseng	tserng	tseeng	tsenq
cha	cha	ch'a	ㄔㄚ	cha	char	chaa	chah
chai	chai	ch'ai	ㄔㄞ	chai	chair	chae	chay
chan	chan	ch'an	ㄔㄢ	chan	charn	chaan	chann
chang	chang	ch'ang	ㄔㄤ	chang	charng	chaang	chanq
chao	chau	ch'ao	ㄔㄠ	chau	chaur	chao	chaw
che	che	ch'e	ㄔㄜ	che	cher	chee	cheh
chen	chen	ch'en	ㄔㄣ	chen	chern	cheen	chenn
cheng	cheng	ch'eng	ㄔㄥ	cheng	cherng	cheeng	chenq
chi	chr	ch'ih	ㄔ	chy	chyr	chyy	chyh
chong	chung	ch'ung	ㄔㄨㄥ	chong	chorng	choong	chonq
chou	chou	ch'ou	ㄔㄡ	chou	chour	choou	chow
chu	chu	ch'u	ㄔㄨ	chu	chwu	chuu	chuh
chua	chwa	ch'ua	ㄔㄨㄚ	chua	chwa	choa	chuah
chuai	chwai	ch'uai	ㄔㄨㄞ	chuai	chwai	choai	chuay
chuan	chwan	ch'uan	ㄔㄨㄢ	chuan	chwan	choan	chuann
chuang	chwang	ch'uang	ㄔㄨㄤ	chuang	chwang	choang	chuanq
chui	chwei	ch'ui	ㄔㄨㄟ	chuei	chwei	choei	chuey
chun	chwun	ch'un	ㄔㄨㄣ	chuen	chwen	choen	chuenn
chuo	chwo	ch'o	ㄔㄨㄛ	chuo	chwo	chuoo	chuoh
ci	tsz	tz'u	ㄘ	tsy	tsyr	tsyy	tsyh
cong	tsung	ts	ㄘㄨㄥ	tsong	tsorng	tsoong	tsonq
cou	tsou	ts'ou	ㄘㄡ	tsou	tsour	tsoou	tsow
cu	tsu	ts'u	ㄘㄨ	tsu	tswu	tsuu	tsuh
cuan	tswan	ts'uan	ㄘㄨㄢ	tsuan	tswan	tsoan	tsuann
cui	tswei	ts'ui	ㄘㄨㄟ	tsuei	tswei	tsoei	tsuey
cun	tswun	ts'un	ㄘㄨㄣ	tsuen	tswen	tsoen	tsuenn
cuo	tswo	ts'o	ㄘㄨㄛ	tsuo	tswo	tsuoo	tsuoh
da	da	ta	ㄉㄚ	da	dar	daa	dah
dai	dai	tai	ㄉㄞ	dai	dair	dae	day
dan	dan	tan	ㄉㄢ	dan	darn	daan	dann

PY	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
dang	dang	tang	ㄉㄤ	dang	darng	daang	danq
dao	dau	tao	ㄉㄠ	dau	daur	dao	daw
de	de	te	ㄉㄜ	de	der	dee	deh
dei	dei	tei	ㄉㄟ	dei	deir	deei	dey
deng	deng	teng	ㄉㄥ	deng	derng	deeng	denq
di	di	ti	ㄉㄧ	di	dyi	dii	dih
dian	dyan	tien	ㄉㄧㄢ	dian	dyan	dean	diann
diao	dyau	tiao	ㄉㄧㄠ	diau	dyau	deau	diaw
die	dye	tieh	ㄉㄧㄝ	die	dye	diee	dieh
ding	ding	ting	ㄉㄧㄥ	ding	dyng	diing	dinq
diu	dyou	tiu	ㄉㄧㄡ	diou	dyou	deou	diow
dong	dung	tung	ㄉㄨㄥ	dong	dorng	doong	donq
dou	dou	tou	ㄉㄡ	dou	dour	doou	dow
du	du	tu	ㄉㄨ	du	dwu	duu	duh
duan	dwan	tuan	ㄉㄨㄢ	duan	dwan	doan	duann
dui	dwei	tui	ㄉㄨㄟ	duei	dwei	doei	duey
dun	dwun	tun	ㄉㄨㄣ	duen	dwen	doen	duenn
duo	dwo	to	ㄉㄨㄛ	duo	dwo	duoo	duoh
e	e	e,o	ㄜ, ㄛ	e	er	ee	eh
ei	ei	ei	ㄟ	ei	eir	eei	ey
en	en	en	ㄣ	en	ern	een	enn
eng	eng	eng	ㄥ	eng	erng	eeng	enq
er	er	erh	ㄦ	erl	erl	eel	ell
fa	fa	fa	ㄈㄚ	fa	far	faa	fah
fan	fan	fan	ㄈㄢ	fan	farn	faan	fann
fang	fang	fang	ㄈㄤ	fang	farng	faang	fanq
fei	fei	fei	ㄈㄟ	fei	feir	feei	fey
fen	fen	fen	ㄈㄣ	fen	fern	feen	fenn
feng	feng	feng	ㄈㄥ	feng	ferng	feeng	fenq
fo	fwo	fo	ㄈㄛ	fo	for	foo	foh
fou	fou	fou	ㄈㄡ	fou	four	foou	fow

PY	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
fu	fu	fu	ㄈㄨ	fu	fwu	fuu	fuh
ga	ga	ka	ㄍㄚ	ga	gar	gaa	gah
gai	gai	kai	ㄍㄞ	gai	gair	gae	gay
gan	gan	kan	ㄍㄢ	gan	garn	gaan	gann
gang	gang	kang	ㄍㄤ	gang	garng	gaang	ganq
gao	gau	kao	ㄍㄠ	gau	gaur	gao	gaw
ge	ge	ke, ko	ㄍㄜ	ge	ger	gee	geh
gei	gei	kei	ㄍㄟ	gei	geir	geei	gey
gen	gen	ken	ㄍㄣ	gen	gern	geen	genn
geng	geng	keng	ㄍㄥ	geng	gerng	geeng	genq
gong	gung	kung	ㄍㄨㄥ	gong	gorng	goong	gonq
gou	gou	kou	ㄍㄡ	gou	gour	goou	gow
gu	gu	ku	ㄍㄨ	gu	gwu	guu	guh
gua	gwa	kua	ㄍㄨㄚ	gua	gwa	goa	guah
guai	gwai	kuai	ㄍㄨㄞ	guai	gwai	goai	guay
guan	gwan	kuan	ㄍㄨㄢ	guan	gwan	goan	guann
guang	gwang	kuang	ㄍㄨㄤ	guang	gwang	goang	guanq
gui	gwei	kuei	ㄍㄨㄟ	guei	gwei	goei	guey
gun	gwun	kun	ㄍㄨㄣ	guen	gwen	goan	guenn
guo	gwo	kuo	ㄍㄨㄛ	guo	gwo	guoo	guoh
ha	ha	ha	ㄏㄚ	ha	har	haa	hah
hai	hai	hai	ㄏㄞ	hai	hair	hae	hay
han	han	han	ㄏㄢ	han	harn	haan	hann
hang	hang	hang	ㄏㄤ	hang	harng	haang	hanq
hao	hau	hao	ㄏㄠ	hau	haur	hao	haw
he	he	he	ㄏㄜ	he	her	hee	heh
hei	hei	hei	ㄏㄟ	hei	heir	heei	hey
hen	hen	hen	ㄏㄣ	hen	hern	heen	henn
heng	heng	heng	ㄏㄥ	heng	herng	heeng	henq
hong	hung	hung	ㄏㄨㄥ	hong	horng	hoong	honq
hou	hou	hou	ㄏㄡ	hou	hour	hoou	how

				ROMATZYH			
[illegible]	YALE	WG	ZYZM	1	2	3	4
hu	hu	hu	ㄏㄨ	hu	hwu	huu	huh
nua	hwa	hua	ㄏㄨㄚ	hua	hwa	hoa	huah
huai	hwai	huai	ㄏㄨㄞ	huai	hwai	hoai	huay
huan	hwan	huan	ㄏㄨㄢ	huan	hwan	hoan	huann
huang	hwang	huang	ㄏㄨㄤ	huang	hwang	hoang	huanq
·ui	hwei	hui	ㄏㄨㄟ	huei	hwei	hoei	huey
hun	hwun	hun	ㄏㄨㄣ	huen	hwen	hoen	huenn
huo	hwo	huo	ㄏㄨㄛ	huo	hwo	huoo	huoh
ji	ji	chi	ㄐㄧ	ji	jyi	jii	jih
jia	jya	chia	ㄐㄧㄚ	jia	jya	jea	jiah
jian	jyan	chien	ㄐㄧㄢ	jian	jyan	jean	jiann
jiang	jyang	chiang	ㄐㄧㄤ	jiang	jyang	jeang	jianq
jiao	jyau	chiao	ㄐㄧㄠ	jiau	jyau	jeau	jiaw
jie	jye	chieh	ㄐㄧㄝ	jie	jye	jiee	jieh
jin	jin	chin	ㄐㄧㄣ	jin	jyn	jiin	jinn
jing	jing	ching	ㄐㄧㄥ	jing	jyng	jiing	jinq
jiong	jyung	chiung	ㄐㄩㄥ	jiong	jyong	jeong	jionq
jiu	jyou	chiu	ㄐㄧㄡ	jiou	jyou	jeou	jiow
ju	jyu	chü	ㄐㄩ	jiu	jyu	jeu	jiuh
juan	jywan	chüan	ㄐㄩㄢ	jiuan	jyuan	jeuan	jiuann
jue	jywe	chüeh	ㄐㄩㄝ	jiue	jyue	jeue	jiueh
jun	jyun	chün	ㄐㄩㄣ	jiun	jyun	jeun	jiunn
ka	ka	k'a	ㄎㄚ	ka	kar	kaa	kah
kai	kai	k'ai	ㄎㄞ	kai	kair	kae	kay
kan	kan	k'an	ㄎㄢ	kan	karn	kaan	kann
kang	kang	k'ang	ㄎㄤ	kang	karng	kaang	kanq
kao	kau	k'ao	ㄎㄠ	kau	kaur	kao	kaw
ke	ke	k'e,k'o	ㄎㄜ	ke	ker	kee	keh
ken	ken	k'en	ㄎㄣ	ken	kern	keen	kenn
keng	keng	k'eng	ㄎㄥ	keng	kerng	keenq	kenq
kong	kung	k'ung	ㄎㄨㄥ	kong	korng	koong	konq
kou	kou	k'ou	ㄎㄡ	kou	kour	koou	kow

PY	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
ku	ku	k'u	ㄎㄨ	ku	kwu	kuu	kuh
kua	kwa	k'ua	ㄎㄨㄚ	kua	kwa	koa	kuah
kuai	kwai	k'uai	ㄎㄨㄞ	kuai	kwai	koai	kuay
kuan	kwan	k'uan	ㄎㄨㄢ	kuan	kwan	koan	kuann
kuang	kwang	k'uang	ㄎㄨㄤ	kuang	kwang	koang	kuanq
kui	kwei	k'uei	ㄎㄨㄟ	kuei	kwei	koei	kuey
kun	kwun	k'un	ㄎㄨㄣ	kuen	kwen	koen	kuenn
kuo	kwo	k'uo	ㄎㄨㄛ	kuo	kwo	kuoo	kuoh
la	la	la	ㄌㄚ	lha	la	laa	lah
lai	lai	lai	ㄌㄞ	lhai	lai	lae	lay
lan	lan	lan	ㄌㄢ	lhan	lan	laan	lann
lang	lang	lang	ㄌㄤ	lhang	lang	laang	lanq
lao	lau	lao	ㄌㄠ	lhau	lau	lao	law
le	le	le	ㄌㄜ	lhe	le	lee	leh
lei	lei	lei	ㄌㄟ	lhei	lei	leei	lei
leng	leng	leng	ㄌㄥ	lheng	leng	leeng	lenq
li	li	li	ㄌㄧ	lhi	li	lii	lih
lia	lya	lia	ㄌㄧㄚ	lhia	lia	lea	liah
lian	lyan	lien	ㄌㄧㄢ	lhian	lian	lean	liann
liang	lyang	liang	ㄌㄧㄤ	lhiang	liang	leang	lianq
liao	lyau	liao	ㄌㄧㄠ	lhiau	liau	leau	liaw
lie	lye	lieh	ㄌㄧㄝ	lhie	lie	liee	lieh
lin	lin	lin	ㄌㄧㄣ	lhin	lin	liin	linn
ling	ling	ling	ㄌㄧㄥ	lhing	ling	liing	linq
liu	lyou	liu	ㄌㄧㄡ	lhiou	liou	leou	liow
long	lung	lung	ㄌㄨㄥ	lhong	long	loong	lonq
lou	lou	lou	ㄌㄡ	lhou	lou	loou	low
lu	lu	lu	ㄌㄨ	lhu	lu	luu	luh
luan	lwan	luan	ㄌㄨㄢ	lhuan	luan	loan	luann
lun	lwun	lun, lün	ㄌㄨㄣ	lhuen	luen	loen	luenn
luo	lwo	lo	ㄌㄨㄛ	lhou	luo	luoo	luoh
lü	lyu	lü	ㄌㄩ	lhiu	liu	leu	liuh
luän	lywan	luän	ㄌㄩㄢ	lhiuan	liuan	leuan	liuann

PY	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
lüe	lywe	lüeh	ㄌㄩㄝ	lhue	liue	leue	liueh
lün	lyun	lün	ㄌㄩㄣ	lhiun	liun	leun	liunh
ma	ma	ma	ㄇㄚ	mha	ma	maa	mah
mai	mai	mai	ㄇㄞ	mhai	mai	mae	may
man	man	man	ㄇㄢ	mhan	man	maan	mann
mang	mang	mang	ㄇㄤ	mhang	mang	maang	manq
mao	mau	mao	ㄇㄠ	mhau	mau	mao	maw
mei	mei	mei	ㄇㄟ	mhei	mei	meei	mey
men	men	men	ㄇㄣ	mhen	men	meen	menn
meng	meng	meng	ㄇㄥ	mheng	men	meeng	menq
mi	mi	mi	ㄇㄧ	mhi	mi	mii	mih
mian	myan	mien	ㄇㄧㄢ	mhian	mian	mean	miann
miao	myau	miao	ㄇㄧㄠ	mhiau	miau	meau	miew
mie	mye	mieh	ㄇㄧㄝ	mhie	mie	miee	mieh
min	min	min	ㄇㄧㄣ	mhin	min	miin	minn
ming	ming	ming	ㄇㄧㄥ	mhing	ming	miing	minq
miu	myou	miu	ㄇㄧㄡ	mhiou	miou	meou	miow
mo	mwo	mo	ㄇㄛ	mho	mo	moo	moh
mou	mou	mou	ㄇㄡ	mhou	mou	moou	mow
mu	mu	mu	ㄇㄨ	mhu	mu	muu	muh
na	na	na	ㄋㄚ	nha	na	naa	nah
nai	nai	nai	ㄋㄞ	nhai	nai	nae	nay
nan	nan	nan	ㄋㄢ	nhan	nan	naan	nann
nang	nang	nang	ㄋㄤ	nhang	nang	naang	nanq
nao	nau	nao	ㄋㄠ	nhau	nau	nao	naw
ne	ne	ne	ㄋㄜ	nhe	ne	nee	neh
nei	nei	nei	ㄋㄟ	nhei	nei	neei	ney
nen	nen	nen	ㄋㄣ	nhen	nen	neen	nenn
neng	neng	neng	ㄋㄥ	nheng	neng	neeng	nenq
ni	ni	ni	ㄋㄧ	nhi	ni	nii	nih
nian	nyan	nien	ㄋㄧㄢ	nhian	nian	nean	niann

PY	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
niang	nyang	niang	ㄋㄧㄤ	nhiang	niang	neang	nianq
niao	nyau	niao	ㄋㄧㄠ	nhiau	niau	neau	niaw
nie	nye	nieh	ㄋㄧㄝ	nhie	nie	niee	nieh
nin	nin	nin	ㄋㄧㄣ	nhin	nin	niin	ninn
ning	ning	ning	ㄋㄧㄥ	nhing	ning	niing	ninq
niu	nyou	niu	ㄋㄧㄡ	nhiu	niou	neou	niow
nong	nung	nung	ㄋㄨㄥ	nhong	nong	noong	nonq
nou	nou	nou	ㄋㄡ	nhou	nou	noou	now
nu	nu	nu	ㄋㄨ	nhu	nu	nuu	nuh
nuan	nwan	nuan	ㄋㄨㄢ	nhuan	nuan	noan	nuann
nun	nwun	nun	ㄋㄨㄣ	nhuen	nuen	noen	nuenn
nuo	nwo	no	ㄋㄨㄛ	nhuo	nuo	nuoo	nuoh
nü	nyu	nü	ㄋㄩ	nhiu	niu	neu	niuh
nüe	nywe	nüeh	ㄋㄩㄝ	nhiue	niue	neue	niueh
ou	ou	ou	ㄡ	ou	our	oou	ow
pa	pa	p'a	ㄆㄚ	pa	par	paa	pah
pai	pai	p'ai	ㄆㄞ	pai	pair	pae	pay
pan	pan	p'an	ㄆㄢ	pan	parn	paan	pann
pang	pang	p'ang	ㄆㄤ	pang	parng	paang	panq
pao	pau	p'ao	ㄆㄠ	pau	paur	pao	paw
pei	pei	p'ei	ㄆㄟ	pei	peir	peei	pey
pen	pen	p'en	ㄆㄣ	pen	pern	peen	penn
peng	peng	p'eng	ㄆㄥ	peng	perng	peeng	penq
pi	pi	p'i	ㄆㄧ	pi	pyi	pii	pih
pian	pyan	p'ien	ㄆㄧㄢ	pian	pyan	pean	piann
piao	pyau	p'iao	ㄆㄧㄠ	piau	pyau	peau	piaw
pie	pye	p'ieh	ㄆㄧㄝ	pie	pye	piee	pieh
pin	pin	p'in	ㄆㄧㄣ	pin	pyn	piin	pinn
ping	ping	p'ing	ㄆㄧㄥ	ping	pyng	piing	pinq
po	pwo	p'o	ㄆㄛ	po	por	poo	poh
pou	pou	p'ou	ㄆㄡ	pou	pour	poou	pow

Pí	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
pu	pu	p'u	ㄆㄨ	pu	pwu	puu	puh
qi	chi	ch'i	ㄑㄧ	chi	chyi	chii	chih
qia	chya	ch'ia	ㄑㄧㄚ	chia	chya	chea	chiah
qian	chyan	ch'ien	ㄑㄧㄢ	chian	chyan	chean	chiann
qiang	chyang	ch'iang	ㄑㄧㄤ	chiang	chyang	cheang	chianq
qiao	chyau	ch'iao	ㄑㄧㄠ	chiau	chyau	cheau	chiaw
qie	chye	ch'ieh	ㄑㄧㄝ	chie	chye	chiee	chieh
qin	chin	ch'in	ㄑㄧㄣ	chin	chyn	chiin	chinn
qing	ching	ch'ing	ㄑㄧㄥ	ching	chyng	chiing	chinq
qiong	chyung	ch'iung	ㄑㄩㄥ	chiong	chyong	cheong	chionq
qiu	chyou	ch'iu	ㄑㄧㄡ	chiou	chyou	cheou	chiow
qu	chyu	ch'ü	ㄑㄩ	chiu	chyu	cheu	chiuh
quan	chywan	ch'üan	ㄑㄩㄢ	chiuan	chyuan	cheuan	chiuann
que	chywe	ch'üeh	ㄑㄩㄝ	chiue	chyue	cheue	chiueh
qun	chyun	ch'ün	ㄑㄩㄣ	chiun	chyun	cheun	chiunn
ran	ran	jan	ㄖㄢ	rhan	ran	raan	rann
rang	rang	jang	ㄖㄤ	rhang	rang	raang	ranq
rao	rau	jao	ㄖㄠ	rhau	rau	rao	raw
re	re	je	ㄖㄜ	rhe	re	ree	reh
ren	ren	jen	ㄖㄣ	rhen	ren	reen	renn
reng	reng	jeng	ㄖㄥ	rheng	reng	reeng	renq
ri	r	jih	ㄖ	rhy	ry	ryy	ryh
rong	rung	jung	ㄖㄨㄥ	rhong	rong	roong	ronq
rou	rou	jou	ㄖㄡ	rhou	rou	roou	row
ru	ru	ju	ㄖㄨ	rhu	ru	ruu	ruh
ruan	rwan	juan	ㄖㄨㄢ	rhuan	ruan	roan	ruann
rui	rwei	jui	ㄖㄨㄟ	rhuei	ruei	roei	ruey
run	rwun	jun	ㄖㄨㄣ	rhuen	ruen	roen	ruann
ruo	rwo	jo	ㄖㄨㄛ	rhuo	ruo	ruoo	ruoh
sa	sa	sa	ㄙㄚ	sa	sar	saa	sah

				ROMATZYH			
PY	YALE	WG	ZYZM	1	2	3	4
sai	sai	sai	ㄙㄞ	sai	sair	sae	say
san	san	san	ㄙㄢ	san	sarn	saan	sann
sang	sang	sang	ㄙㄤ	sang	sarng	saang	sanq
sao	sau	sao	ㄙㄠ	sau	saur	sao	saw
se	se	se	ㄙㄜ	se	ser	see	seh
sen	sen	sen	ㄙㄣ	sen	sern	seen	senn
seng	seng	seng	ㄙㄥ	seng	serng	seeng	senq
sha	sha	sha	ㄕㄚ	sha	shar	shaa	shah
shai	shai	shai	ㄕㄞ	shai	shair	shae	shay
shan	shan	shan	ㄕㄢ	shan	sharn	shaan	shann
shang	shang	shang	ㄕㄤ	shang	sharng	shaang	shanq
shao	shau	shao	ㄕㄠ	shau	shaur	shao	shaw
she	she	she	ㄕㄜ	she	sher	shee	sheh
shei	shei	shei	ㄕㄟ	shei	sheir	sheei	shey
shen	shen	shen	ㄕㄣ	shen	shern	sheen	shenn
sheng	sheng	sheng	ㄕㄥ	sheng	sherng	sheeng	shenq
shi	shr	shih	ㄕ	shy	shyr	shyy	shyh
shou	shou	shou	ㄕㄡ	shou	shour	shoou	show
shu	shu	shu	ㄕㄨ	shu	shwu	shuu	shuh
shua	shwa	shua	ㄕㄨㄚ	shua	shwa	shoa	shuah
shuai	shwai	shuai	ㄕㄨㄞ	shuai	shwai	shoai	shuay
shuan	shwan	shuan	ㄕㄨㄢ	shuan	shwan	shoan	shuann
shuang	shwang	shuang	ㄕㄨㄤ	shuang	shwang	shoang	shuanq
shui	shwei	shui	ㄕㄨㄟ	shuei	shwei	shoei	shuey
shun	shwun	shun	ㄕㄨㄣ	shuen	shwen	shoen	shuenn
shuo	shwo	shuo	ㄕㄨㄛ	shuo	shwo	shuoo	shuoh
si	sz	szu, ssu	ㄙ	sy	syr	syy	syh
song	sung	sung	ㄙㄨㄥ	song	sorng	soong	sonq
sou	sou	sou	ㄙㄡ	sou	sour	soou	sow
su	su	su	ㄙㄨ	su	swu	suu	suh
suan	swan	suan	ㄙㄨㄢ	suan	swan	soan	suann
sui	swei	sui	ㄙㄨㄟ	suei	swei	soei	suey
sun	swun	sun	ㄙㄨㄣ	suen	swen	soen	suenn
suo	swo	so	ㄙㄨㄛ	suo	swo	suoo	suoh

PY	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
ta	ta	t'a	ㄊㄚ	ta	tar	taa	tah
tai	tai	t'ai	ㄊㄞ	tai	tair	tae	tay
tan	tan	t'an	ㄊㄢ	tan	tarn	taan	tann
tang	tang	t'ang	ㄊㄤ	tang	tarng	taang	tanq
tao	tau	t'ao	ㄊㄠ	tau	taur	tao	taw
te	te	t'e	ㄊㄜ	tc	ter	tee	teh
teng	teng	t'eng	ㄊㄥ	teng	terng	teeng	tenq
ti	ti	t'i	ㄊㄧ	ti	tyi	tii	tih
tian	tyan	t'ien	ㄊㄧㄢ	tian	tyan	tean	tiann
tiao	tyau	t'iao	ㄊㄧㄠ	tiau	tyau	teau	tiaw
tie	tye	t'ieh	ㄊㄧㄝ	tie	tye	tiee	tieh
ting	ting	t'ing	ㄊㄧㄥ	ting	tyng	tiing	tinq
tong	tung	t'ung	ㄊㄨㄥ	tong	torng	toong	tonq
tou	tou	t'ou	ㄊㄡ	tou	tour	toou	tow
tu	tu	t'u	ㄊㄨ	tu	twu	tuu	tuh
tuan	twan	t'uan	ㄊㄨㄢ	tuan	twan	toan	tuann
tui	twei	t'ui	ㄊㄨㄟ	tuei	twei	tooi	tuey
tun	twun	t'un	ㄊㄨㄣ	tuen	twen	toen	tuenn
tuo	two	t'o	ㄊㄨㄛ	tuo	two	tuoo	tuoh
wa	wa	wa	ㄨㄚ	ua	wa	woa	wah
wai	wai	wai	ㄨㄞ	uai	wai	woai	way
wan	wan	wan	ㄨㄢ	uan	wan	woan	wann
wang	wang	wang	ㄨㄤ	uang	wang	woang	wanq
wei	wei	wei	ㄨㄟ	uei	wei	woei	wey
wen	wen	wen	ㄨㄣ	uen	wen	woen	wenn
weng	weng	weng	ㄨㄥ	ueng	weng	woeng	weng
wo	wo	wo	ㄨㄛ	uo	wo	woo	woh
wu	wu	wu	ㄨ	u	wu	wuu	wuh
xi	syi	hsi	ㄒㄧ	shi	shyi	shii	shih
xia	sya	hsia	ㄒㄧㄚ	shia	shya	shea	shiah
xian	syan	hsien	ㄒㄧㄢ	shian	shyan	shean	shiann

PY	YALE	WG	ZYZM	ROMATZYH 1	2	3	4
xiang	syang	hsiang	ㄒㄧㄤ	shiang	shyang	sheang	shianq
xiao	syau	hsiao	ㄒㄧㄠ	shiau	shyau	sheau	shiaw
xie	sye	hsieh	ㄒㄧㄝ	shie	shye	shiee	shieh
xin	syin	hsin	ㄒㄧㄣ	shin	shyn	shiin	shinn
xing	sying	hsing	ㄒㄧㄥ	shing	shyng	shiing	shinq
xiong	syung	hsiung	ㄒㄩㄥ	shiong	shyong	sheong	shionq
xiu	syou	shiu	ㄒㄧㄡ	shiou	shyou	sheou	shiow
xu	syu	shü	ㄒㄩ	shiu	shyu	sheu	shiuh
xuan	sywan	hsüan	ㄒㄩㄢ	shiuan	shyuan	sheuan	shiuann
xue	sywe	shüeh	ㄒㄩㄝ	shiue	shyue	sheue	shiueh
xun	syun	hsün	ㄒㄩㄣ	shiun	shyun	sheun	shiunn
ya	ya	ya	ㄧㄚ	ia	ya	yaa	yah
yai	yai	yai	ㄧㄞ	iai	yai	yae	yay
yan	yan	yen	ㄧㄢ	ian	yan	yean	yann
yang	yang	yang	ㄧㄤ	iang	yang	yeang	yanq
yao	yau	yao	ㄧㄠ	iau	yau	yeau	yaw
ye	ye	yeh	ㄧㄝ	ie	ye	yee	yeh
yi	yi	i	ㄧ	i	yi	yi	ih
yin	yin	yin	ㄧㄣ	in	yn	yiin	yinn
ying	ying	ying	ㄧㄥ	ing	yng	yiing	yinq
yong	yung	yung	ㄩㄥ	iong	yong	yeong	yonq
you	you	yu	ㄧㄡ	iou	you	yeou	yow
yu	yu	yü	ㄩ	iu	yu	yeu	yuh
yuan	ywan	yüan	ㄩㄢ	iuan	yuan	yeuan	yuann
yue	ywe	yüeh	ㄩㄝ	iue	yue	yeue	yueh
yun	yun	yün	ㄩㄣ	iun	yun	yeun	yunn
za	dza	tsa	ㄗㄚ	tza	tzar	tzaa	tzah
zai	dzai	tsai	ㄗㄞ	tzai	tzair	tzae	tzay
zan	dzan	tsan	ㄗㄢ	tzan	tzarn	tzaan	tzann
zang	dzang	tsang	ㄗㄤ	tzang	tzarng	tzaang	tzanq
zao	dzau	tsao	ㄗㄠ	tzau	tzaur	tzao	tzaw

				ROMATZYH			
PY	YALE	WG	ZYZM	1	2	3	4
ze	dze	tse	ㄗㄜ	tze	tzer	tzee	tzeh
zei	dzei	tsei	ㄗㄟ	tzei	tzeir	tzeei	tzey
zen	dzen	tsen	ㄗㄣ	tzen	tzern	tzeen	tzenn
zeng	dzeng	tseng	ㄗㄥ	tzeng	tzerng	tzeeng	tzenq
zha	ja	cha	ㄓㄚ	ja	jar	jaa	jah
zhai	jai	chai	ㄓㄞ	jai	jair	jae	jay
zhan	jan	chan	ㄓㄢ	jan	jarn	jaan	jann
zhang	jang	chang	ㄓㄤ	jang	jarng	jaang	janq
zhao	jau	chao	ㄓㄠ	jau	jaur	jao	jaw
zhe	je	che	ㄓㄜ	je	jer	jee	jeh
zhei	jei	chei	ㄓㄟ	jei	jeir	jeei	jey
zhen	jen	chen	ㄓㄣ	jen	jern	jeen	jenn
zheng	jeng	cheng	ㄓㄥ	jeng	jerng	jeeng	jenq
zhi	jr	chih	ㄓ	jy	jyr	jyy	jyh
zhong	jung	chung	ㄓㄨㄥ	jong	jorng	joong	jonq
zhou	jou	chou	ㄓㄡ	jou	jour	joou	jow
zhu	ju	chu	ㄓㄨ	ju	jwu	juu	juh
zhua	jwa	chua	ㄓㄨㄚ	jua	jwa	joa	juah
zhuai	jwai	chuai	ㄓㄨㄞ	juai	jwai	joai	juay
zhuan	jwan	chuan	ㄓㄨㄢ	juan	jwan	joan	juann
zhuang	jwang	chuang	ㄓㄨㄤ	juang	jwang	joang	juanq
zhui	jwei	chui	ㄓㄨㄟ	jaei	jwei	joei	juey
zhun	jwun	chun	ㄓㄨㄣ	juen	jwen	joen	juenn
zhuo	jwo	cho	ㄓㄨㄛ	juo	jwo	juoo	juoh
zi	dz	tzu	ㄗ	tzy	tzyr	tzyy	tzyh
zong	dzung	tsung	ㄗㄨㄥ	tzong	tzorng	tzoong	tzonq
zou	dzou	tsou	ㄗㄡ	tzou	tzour	tzoou	tzow
zu	dzu	tsu	ㄗㄨ	tzu	tzwu	tzuu	tzuh
zuan	dzwan	tsuan	ㄗㄨㄢ	tzuan	tzwan	tzoa..	tzuann
zui	dzwei	tsui	ㄗㄨㄟ	tzuei	tzwei	tzoei	tzuey
zun	dzwun	tsun	ㄗㄨㄣ	tzuen	tzwen	tzoen	tzuenn
zuo	dzwo	tso	ㄗㄨㄛ	tzuo	tzwo	tzuoo	tzuoh

CONVERSION ALGORITHMS AMONG DIFFERENT TRANSCRIPTIONS

K. P. Li

I. Introduction:

A set of conversion rules for the Chinese Transcription Systems Pinyin, Yale, Wade-Giles and Romatzyh has been constructed in the form of regular linguistic rules which can be applied to computer programs. The conditions and stored information are reduced to the minimum, including only the alphabets used in the systems, the tone marks and word boundaries. The traffic rules used here are very simple, as shown in Fig. 1. In each step, the TEST is to find out whether any portion of an input is exactly the same as that on the left hand side of the function; if the response is YES, then the said portion is changed to the part shown in the right hand side of the function; if a NO is indicated, then the input goes to the next step without change. The same process is carried through until the END is reached. During the TEST a set of elements may be listed in one single step, in which case the input must be checked against each element in the order listed. Once an element is found to be identical with a portion of the input, the same conversion operation takes place, after which the input must be carried directly into the next step without checking the remaining elements in the list. In other words, each step contains no more than one operation. A complete program of conversion rules between Pinyin-Yale, Pinyin-WG, and Pinyin-Romatzyh (Basic form), in both directions, is listed in Section II.

Before applying the rules, it is essential to have the word boundaries (•) marked out. Unfortunately, all systems contain some ambiguity in word boundary which necessitates pre-marking by hand. In Section III we have constructed a minimum set of rules for marking out by hand the ambiguous word boundaries and a program for indicating all word boundaries which can be done by machine, in the Pinyin System. Similar procedures can be followed for constructing rules and programs for the other systems but they are much more complicated than those for the Pinyin System.

The tonal spelling in the Romatzyh System (shown in Section IV) presents additional complications in programming. As a result, more complex traffic rules are felt necessary in order to reduce the considerable number of conversion rules. The flow chart for these traffic rules is shown in Fig. 2. We have not yet investigated the possibility of simplifying the rules in Section II by the application of these traffic rules.

In Section V, several examples illustrating the actual applications of these rules are given. We intend to run a computer program to check these rules in the future. The references upon which the various transcription systems are based are listed in Section VI.

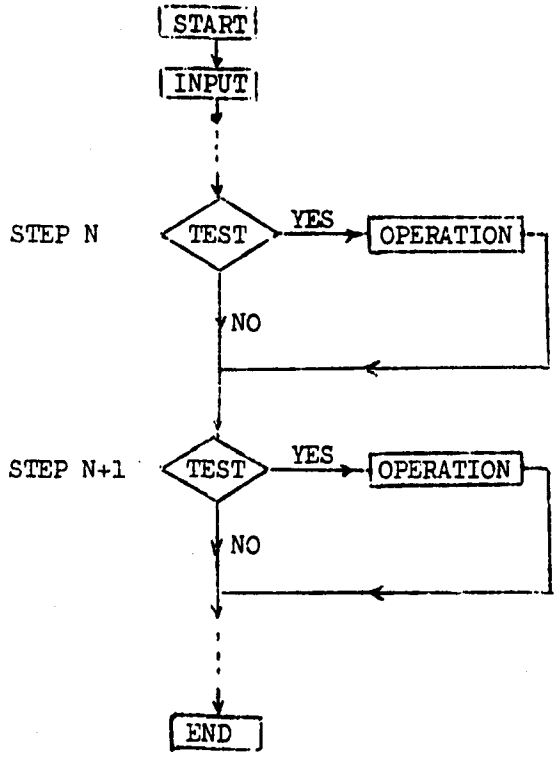


Figure 1. The pattern of traffic rules used in the Conversion Rules between systems. The output of Step N always goes to the next step.

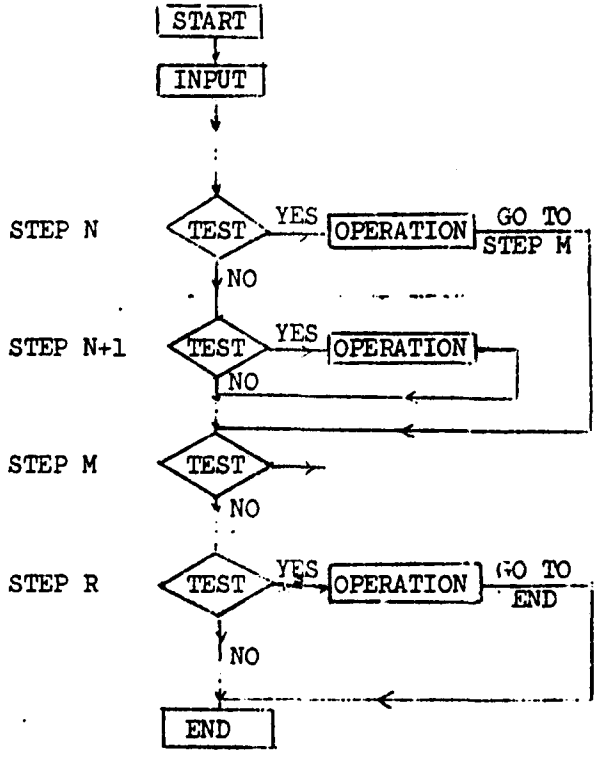


Figure 2. Three different traffic rules used in tonal spelling. The output of Step N may go to the next step, or to some later Step M, or to the end.

II. Conversion Rules among Chinese Transcription Systems:

A. Conversion Rules from Pinyin to Yale:

X_1: .b .p .m .f

X_2: i*d. u*i. a*o.

X_3: io*u. ue*i. au*.

* is tone mark.

X_4: .j .q .x

X_5: .d .t .m .l .g .k .h

X_6: .zh .ch .sh .r

X_7: .z .c .s

X_8: X_5 X_6 X_7

X_9: e a o u

X_{10}: i u

X_{11}: y w

X_{12}: .zh .q .x .z .ch .c

X_{13}: .j .ch .sy .dz .ch .ts

Rules:

1. $X_1\overset{*}{e}.$ ⟶ $X_1u\overset{*}{e}.$
2. X_2 ⟶ X_3
3. X_4u or $X_4\overset{*}{u}$ ⟶ X_4i or $X_4\overset{*}{u}$
4. $X_8\overset{*}{u}n$ ⟶ $X_8u\overset{*}{u}n$
5. ü or $\overset{*}{u}$ ⟶ yu or $y\overset{*}{u}$
6. $X_{10}\overset{*}{X}_9$ ⟶ $X_{11}\overset{*}{X}_9$
7. $X_6\overset{*}{i}.$ ⟶ $X_6\overset{*}{r}.$
8. $X_7\overset{*}{i}.$ ⟶ $X_7\overset{*}{z}.$
9. ong. ⟶ ung.
10. X_{12} ⟶ X_{13}
11. XX or $X\overset{*}{X}$ ⟶ X or $\overset{*}{X}$

WRITE OUTPUT.

B. Conversion Rules from Yale to Pinyin.

X_1: .$d\overset{*}{z}$. .dz .ts .j

X_2: .$z\overset{*}{i}$. .z .c .sh

X_3: .zh .ch .s

X_4: .j .q .x

X_5: yu $y\overset{*}{u}$ y $\overset{*}{i}$

X_6: u $\overset{*}{u}$ i $\overset{*}{i}$

X_7: n l

X_8: $\overset{*}{z}$. $\overset{*}{r}$.

X_9: iou. uei. au. .i.
X_{10}: iu. ui. ao. .ri.
X_{11}: .b .p .m .f

Rules:
1. $X_1 \longrightarrow X_2$
2. ung. $\longrightarrow$ ong.
3. w $\longrightarrow$ u
4. .u $\longrightarrow$ .w
5. $X_3X_5 \longrightarrow X_4X_6$
6. X_7 yu or yu $\longrightarrow$ X_7 ü or ü
7. y $\longrightarrow$ i
8. .i $\longrightarrow$ .y
9. $X_8 \longrightarrow$ i.
10. XX or XX $\longrightarrow$ X or X
11. $X_9 \longrightarrow X_{10}$
12. X_{11} uo $\longrightarrow$ X_{11} o

WRITE OUTPUT

C. Conversion Rules from Pinyin to Wade-Giles:

X_1: .j .q .x .y
X_2: .z .c .s
X_3: .zh .ch .sh .r
X_4: .b .p .d .t .g .k .zh .ch .z .c .r .j .q .x
X_5: .p .p' .t .t' .k .k' .ch .ch' .ts .ts' .j .ch .ch' .hs
X_6: .tsz .ts'z
X_7: .tz .tz'
X_8: y i ü
X_9: ian. ong. uo. .er.
X_{10}: ien. ung. o. .erh.
X_{11}: .k .k' .h
X_{12}: o. ui. e.
X_{13}: uo. uei. o.

Rules:
1. X_1u or ů ⟶ X_1U or Ů
2. X_2i. ⟶ X_2zů.
3. X_3i. ⟶ X_3ih.
4. X_4 ⟶ X_5
5. X_6 ⟶ X_7
6. X_8ě. ⟶ X_8ěh.
7. X_9 ⟶ X_{10}
8. $X_{11}X_{12}$ ⟶ $X_{11}X_{13}$

WRITE OUTPUT.

D. Conversion Rules from Wade-Giles to Pinyin:

X_1: zů. .tz'ů. .ssu.
X_2: sǐ. .ts'ǐ. .ssi.
X_3: .ch' .ch .j .ts' .ts .t' .t .p' .p .k' .k
X_4: .ch .zh .r .c .z .t .d .p .b .k .g
X_5: .g .k .h
X_6: .zh .ch .hs .y
X_7: .j .q .x .y
X_8: i ū ǐ Ů
X_9: i u ǐ ů
X_{10}: .ss ř. ůng. h. .yu.
X_{11}: .s ǐ. ong. . .you.
X_{12}: .d .t .n .l .zh .ch .sh .z .c .s .r.
X_{13}: i .y

Rules:
1. X_1 ⟶ X_2
2. X_3 ⟶ X_4
3. ǐh. ⟶ ř.
4. X_5ǒ. ⟶ X_5ě.
5. X_6X_8 ⟶ X_7X_8 ⟶ X_7X_9
6. X_{10} ⟶ X_{11}
7. X_{12}ǒ. ⟶ X_{12}uǒ.
8. X_5uěi. ⟶ X_5uǐ.
9. X_{13}en. ⟶ X_{13}an.

WRITE OUTPUT.

E Conversion Rules from Pinyin to Romatzyh (Basic Form):

X_1: .j .q .x .y
X_2: y w
X_3: i u
X_4: ui. iu. ao.
X_5: uei. iou. au.
X_6: .zh .ch .sh .z .c .s .r
X_7: .zh .ch .c .z .q .x .er.
X_8: .j .ch .ts .tz .ch .sh .el.

Rules:
1. (Input) $\longrightarrow$ Input + Tone
2. $X_1u \longrightarrow X_1ü$
3. un $\longrightarrow$ uen
4. ü $\longrightarrow$ iu
5. $.X_2 \longrightarrow .X_3$
6. .XX $\longrightarrow$ .X
7. $X_6i. \longrightarrow X_6y.$
8. $X_4 \longrightarrow X_5$
9. $X_7 \longrightarrow X_8$

WRITE OUTPUT.

F. Conversion Rules from Romatzyh(Basic Form)to Pinyin:

X_1: .ch .sh .j
X_2: .q .x .v
X_3: .j .tz .ts
X_4: .zh .z .c
X_5: .in .i. .u.
X_6: .iin .ii. .uu.
X_7: .i .u
X_8: .y .w
X_9: uen. uei. iou. au.
X_{10}: un. ui. iu. ao.

Rules:
1. $X_1i \longrightarrow X_2j$
2. $X_3 \longrightarrow X_4$
3. $X_2iu \longrightarrow X_2u$

4. .v ⟶ .j
5. y. ⟶ i.
6. X_5 ⟶ X_6
7. X_7 ⟶ X_8
8. X_9 ⟶ X_{10}
9. Input + Tone ⟶ (Input)#

WRITE OUTPUT.

#Tone mark is always placed on the first vowel except when combined vowels have i, u, or ü as the first vowel, in which case it should be placed on the second vowel.

III. Rules and Program Marking out Word Boundary for Pinyin System:

A. Rule for Eliminating Ambiguous Word Boundaries:

a, e and o not preceded by a consonant must have word boundary (•) marked out by hand.

B. Program for Marking out All Unambiguous Word Boundaries:

X_1: b p m f d t n l g k h z c s r j q x y w and (space)

X_2: z c s

X_3: n r

Rules:
1. X_1 ⟶ $.X_1$
2. $.X_2.h$ ⟶ $.X_2h$
3. $.X_3.$ ⟶ $X_3.$
4. .g. ⟶ g.

C. Examples*

a. 這裏 有 十二把 椅子
zhèlǐ yǒu shíèrbǎ yǐzi.

zhèlǐ yǒu shí.èrbǎ yǐzi.

1. .z.hè.lǐ. .yǒu. .s.hí.è.r.bǎ. .yǐ.zi..

2. .zhè.lǐ. .yǒu. .shí.è.r.bǎ. .yǐ.zi..

*Unnumbered rules are for hand-marked boundaries. Numbered rules are for machine-marked boundaries.

3. .zhè.lǐ. .yǒu. .shí.èr.bǎ. .yǐ.zi..

4. .zhè.li. .yǒu. .shi.èr.bǎ. .yi.zi..

b. 中俄關係 還是 沒有 改變
zhōngèguānxi huánshi méiyǒu gǎibiàn.

zhōng.èguānxi huánshi méiyǒu gǎibiàn.

1. .zh.hō.n.g.è.guā.n.xi. .huá.n.s.hi. .méi.yǒu. .gǎi.biá.n..
2. .zhō.n.g.è.guā.n.xi. .huá.n.shi. .méi.yǒu. .gǎi.bia.n..
3. .zhōn.g.è.guān.xi. .huán.shi. .méi.yǒu. .gǎi.bian..
4. .zhōng.è.guān.xi. .huán.shi. .méi.yǒu. .gǎi.bian..

c. 不按照規則推行 一定會出錯
búànzhàoguīzétuīyǎn yīdìnghuìchūcuò.

bú.ànzhàoguīzétuīyǎn yīdìnghuìchūcuò.

1. .bú.à.n.z.hào.guī.zé.tuī.yǎ.n. .yī.dì.n.g.hui.c.hū.cuò..
2. .bú.à.n.zhào.guī.zé.tuī.yǎ.n. .yī.dì.n.g.hui.chū.cuò..
3. .bú.àn.zhào.guī.zē.tuī.yǎn. .yī.dìn.g.hui.chū.cuò..
4. .bú.àn.zhào.guī.zē.tuī.yǎn. .yī.dìng.hui.chū.cuò..

d.

年糕	男兒	木耳	身高
niángāo	nánér	mùěr	shēngāo
niángāo	nán.ér	mù.ěr	shēngāo
1. .nía.n.gāo.	1. .ná.n.é.r.	1. mù.ě.r.	1. .s.hē.n.gāo.
2. .nía.n.gāo.	2. .ná.n.é.r.	2. .mù.ě.r.	2. .shē.n.gāo.
3. .nían.gāo.	3. .nán.ér.	3. .mù.ěr.	3. .shēn.gāo.
4. .nían.gāo.	4. .nán.ér.	4. .mù.ěr.	4. .shēn.gāo.

IV. Tonal Spelling Rules for Romatzyh System:

A. Conversion Rules from Basic Form to Tonal Spelling:

a. Pre-rules:

X_1: m n l r

1. .X_1 + (1) ⟶ .X_1h go to END
2. .X_1 + (2) ⟶ .X_1 go to END
3. (input) + (1) ⟶ (input) go to END
4. (input) + (2) ⟶ (input) go to B-2 Rules

5. (input) + (3) ⟶ (input) go to B-3 Rules
6. (input) + (4) ⟶ (input) go to B-4 Rules

b. B-2 Rules:

X_1: i u
X_2: e a o y
X_3: i u
X_4: y w
X_5: yi wu
X_6: ay. ey. aw. ow.
X_7: ai. ei. au. ou.

Rules:
1. if X_1 X_2 X_3 go to Step 5
2. X_2X_3 ⟶ X_2X_3r go to OUTPUT
3. if X_1X_2 go to Step 5
4. X_2 ⟶ X_2r go to OUTPUT
5. i ⟶ y
6. u ⟶ w
7. X_6 ⟶ X_7 go to OUTPUT
8. X_4. ⟶ X_5.

WRITE OUTPUT.

c. B-3 Rules:

X_1: a u o
X_2: a e
X_3: i u
X_4: y w
X_5: e o
X_6: ye wo
X_7: a o e y

Rules:
1. iX_1 ⟶ eX_1 go to Step 6
2. uX_2 ⟶ oX_2 go to Step 6
3. aX_3 ⟶ aX_5 go to Step 6
4. X_7 ⟶ X_7X_7 go to Step 6
5. X_3 ⟶ X_3X_3
6. $.X_3X_3$ ⟶ $.X_3X_3X_3$
7. $.X_3$ ⟶ $.X_4$

8. if $.X_7X_7$ go to OUTPUT
9. $.X_5 \longrightarrow .X_6$

WRITE OUTPUT.

d. B-4 Rules:

X_1: i. u. n. ng. l. . (except space before the boundary)
X_2: y. w. nn. nq. ll. h.
X_3: a e o
X_4: y w
X_5: i u

Rules: 1. $X_1 \longrightarrow X_2$
2. if X_3X_4 go to Step 4
3. $X_4. \longrightarrow X_5h.$
4. $.X_5h. \longrightarrow .X_5X_5h.$
5. .in $\longrightarrow$ .iin
6. $.X_5 \longrightarrow .X_4$

WRITE OUTPUT.

B. Conversion Rules from Tonal Spelling to Basic Form:

a. Pre-rules:

X_1: .n .m .l .r
X_2: h. ll. nn. nq.
X_3: y. yy.
X_4: y. w.
X_5: ao ae oa oe ea eo eu
X_6: r y w
X_7: aa ee oo ii uu yy
X_8: .tz .ts .j .s .ch .sh .r

Rules: 1. $X_1h \longrightarrow X_1 + (1)$ go to END
2. if X_2 go to 4-B Rules
3. if X_8X_3 go to Step 5
4. if X_4 go to 4-B Rules
5. if X_5 go to 3-B Rules
6. if X_7 go to 3-B Rules
7. if y. go to Step 10
8. if X_6 go to 2-B Rules

9. $X_1 \longrightarrow X_1$ + (2) go to END
10. (input) $\longrightarrow$ (input) + (1) go to END

b. 2-B Rules:

X_1: , n l
X_2: y w
X_3: i u

Rules: 1. $rX_1 \longrightarrow X_1$
2. if y. go to Step 4
3. $X_2 \longrightarrow X_3$
4. $X_3X_3 \longrightarrow X_3$
5. (input) $\longrightarrow$ (input) + (2)
WRITE OUTPUT.

c. 3-B Rules:

X_1: e o
X_2: i u
X_3: u o
X_4: y w

Rules: 1. $X_1a \longrightarrow X_2a$
2. $aX_1 \longrightarrow aX_2$
3. $eX_3 \longrightarrow iX_3$
4. $X_3e \longrightarrow ue$
5. $.X_4 \longrightarrow .X_2$
6. $XX \longrightarrow X$
7. $XX \longrightarrow X$
8. (input) $\longrightarrow$ (input) + (3)
WRITE OUTPUT.

d. 4-B Rules:

Rules: 1. yh. $\longrightarrow$ y. go to Step 7
2. w $\longrightarrow$ u
3. y $\longrightarrow$ i
4. h. $\longrightarrow$.
5. nq. $\longrightarrow$ ng.
6. $XX \longrightarrow X$

7. (input) ⟶ (input) + (4)
WRITE OUTPUT.

V. Sample Applications of All Rules (Word Boundaries Pre-marked):

A. PY ⟶ Yale

.xuǎn.	.cǔn.	qǐong	.zǐ.
3. .xüǎn.	4. .cuǔn.	6. .qyǒng.	8. .zž.
5. .xyuǎn.	6. .cwǔn.	9. .qyǔng.	10. .dzž.
10. .syyuǎn.	10. .tswǔn.	10. .chyǔng.	11. .dž.
11. .syuǎn.			

B. Yale ⟶ PY

.lywǎn.	.yǎu.	.jyǔng.	chywǎn.
3. .lyuǎn.	7. .iǎu.	1. .zhyǔng.	3. .chyuǎn.
6. .lüǎn.	8. .yǎu.	2. .zhyǒng.	5. .quǎn.
	11. .yǎo.	3. .jiǒng.	

C. PY ⟶ WG

.quě.	.quǒ.	.zǐ.	.zhǒng.
1. .qüě.	4. .kuǒ.	2. .zzǔ.	4. .chǒng.
4. .ch'üě.	7. .kǒ.	4. .tszǔ.	7. .chǔng.
6. .ch'üěh.	8. .kuǒ.	5. .tzǔ.	

D. WG ⟶ PY

.tz'ǔ.	.chiěh.	.ch'ǐh.	.kuěi.
1. .ts'ǐ.	2. .zhiěh.	2. .chǐh.	2. .guěi.
2. .cǐ.	5. .jiěh.	3. .chř.	8. .guǐ.
	6. .jiě.	6. .chǐ.	

E. PY ⟶ RMTZ (Basic)

.qǔn.	.zǔn.	.zhǐ.	.yǐng.
1. .qun. + T	1. .zun. + T	1. .zhi. + T	1. .ying. + T
2. .qün. + T	3. .zuen. + T	7. .zhy. + T	5. .iing. + T

4. .qiun. + T 8. .tzuen. + T 8. .jy. + T 6. .ing. + T
9. .chiun. + T

F. RMTZ (Basic) → PY:

.jau. + T	.jiuan. + T	.iau. + T	.u. + T
2. .zhau. + T	1. .vinàn. + T	7. .yau. + T	6. .uu. + T
8. .zhao. + T	3. .vuan. + T	8. .yao. + T	7. .wu. + T
9. .zhǎo.	4. .juan. + T	9. .yǎo.	9. .wǔ.
	9. .juán.		

G. Tonal Spelling:

1. .tzuei. + (2)
   a. 4. .tzuei. go to b
   b. 1. go to step 5
      5. .tzuey.
      6. .tzwey.
      7. .tzwei. go to OUTPUT

2. .jei. + (2)
   a. 4. .jei. go to b
   b. 2. .jeir. go to OUTPUT

3. .u. + (3)
   a. 5. .u. go to c
   c. 5. .uu.
      6. .uuu.
      7. .wuu.

4. .duei. + (3)
   a. 5. .duei. go to c
   c. 2. .doci. go to step 6

5. .uei. + (4)
   a. 6. .uei. go to d
   d. 1. .uey.
      2. go to step 4
      6. .wey.

6. .mie. + (4)
   a. 6. .mie. go to d
   d. 1. .mieh.

7. .yann
   a. 2. .yann. go to d
   d. 3. .iann.
      6. .ian.
      7. .ian. + (4)

8. .yee.
   a. 6. .yee. go to c
   c. 5. .iee.
      6. .ie.
      8. .ie. + (3)

9. .meau.

a. 5. .meau. go to c

c. 1. .miau.

8. .miau. + (3)

10. .herng.

a. 8. .herng go to b

b. 1. .heng.

5. .heng. + (2)

VI. References for Conversion Table:

Pinyin - Hànyǔ Pinyīn Cíhuì 漢語拼音詞彙 Peiping: Wénzì gǎigé chū bǎnshè, 1958.

Yale - Po-fei Huang, ed. IFEL Vocabulary of Spoken Chinese, New Haven, 1954.

Wade-Giles - Mathews' Chinese-English Dictionary, revised edition, Harvard University Press, 1961.

Zhùyīnzìmǔ - Hànyǔ Cídiǎn 漢語詞典 Peiping: Shāngwù yìnshūguǎn, 1959.

Romatzyh - Yuan-ren Zhao, Mandarin Primer, Cambridge: Harvard University Press, 1948.

NOTE:

We have recently received the Chinese Character Indexes by Ching-yi Doughty, Sydney M. Lamb and Samuel E. Martin, 5 volumes, University of California Press, 1963. In volume I of this work, pp. xix-xxi, rules are given for the automatic conversion between Guoyeu Romatzyh and Pinyin systems. Some similarities were observed between these rules and those in sections II-E, II-F, and IV of this paper. A critical comparison of these sets of rules will be forthcoming.